

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

FACULTAD DE CIENCIAS MATEMÁTICAS

UNIDAD DE POSTGRADO

Media proximal y regularización

TESIS

Para optar el grado académico de Magíster en Matemática Aplicada,
mención Matemática Computacional

AUTOR

Marlo Carranza Purca

ASESOR

Tomás Alberto Núñez Lay

Lima – Perú

2013

Autor:	Marlo Carranza Purca		
Título:	Media proximal y regularización		
Fecha:	17 12 2013	Páginas:	71
Facultad:	UNMSM. Facultad de Ciencias Matemáticas		
Asesor:	Tomás Alberto Núñez Lay		
Palabras	Conjuntos y funciones convexas, transformada convexa		
claves:	Media proximal, Aproximación regular		

MEDIA PROXIMAL Y REGULARIZACIÓN

Marlo Carranza Purca

Tesis presentada a consideración del cuerpo docente de la facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos como parte de los requisitos para obtener el grado académico de Magister en Matemática Aplicada, mención Matemática Computacional.

Aprobada por:

.....
Renato Mario Benazic Tomé

.....
Tomás Alberto Núñez Lay

.....
Luís Enrique Carrillo Díaz

.....
Eladio Teofilo Ocaña Anaya

.....
Victor Rafael Cabanillas Zannini

Dedicatoria

A todos mis seres queridos que me impulsan para seguir avanzando siempre.

A todos mis profesores que me guían con su experiencia y con sus conocimientos

Agradecimientos

Quiero agradecer a Dios por darme la vida y la oportunidad de realizar muchas cosas importantes.

Al profesor Tomas Nuñez Lay, por ser además de mi asesor, un gran maestro y un amigo que siempre me mostró su ayuda, por el invaluable apoyo académico que me brindó, pero sobre todo por la enorme paciencia que me tuvo. Muchísimas gracias profesor.

A los profesores Eladio Ocaña Anaya, Renato Mario Benazic, Luís Enrrique Carrillo Díaz y Victor Rafael Cabanillas Zannini por sus importantes observaciones luego de revisar mi tesis, contribuyendo así a realizar importantes precisiones y así desarrollar de la mejorar manera los capítulos de esta tesis.

A mi amigo Mg.Frank Navarro Rojas, por sus sugerencias y su amistad.

A mi hermano Aldo Carranza Purca, por todo entusiasmo y su apoyo.

A mi esposa y a mi hijo por su tiempo y comprensión .

A los profesores de la Facultad de Ciencias Matemáticas que contribuyeron a mi formación profesional.

Índice general

Introducción	1
1. Análisis convexo	3
1.1. Conjuntos convexos	3
1.2. Función convexa	5
1.2.1. Subgradiente	12
1.2.2. Diferenciabilidad convexa	13
1.2.3. Conjugada de Fenchel y inf-convolución	15
2. La media proximal	31
2.1. Propiedades	34
2.1.1. Conjugada de Fenchel para la media proximal	36
2.1.2. Aproximación regular	41
2.2. Convergencia de la media proximal en presencia de errores de computo	44
3. Aplicación y extensión	56
3.1. Aproximación suave autodual	56
3.1.1. Aplicación a la optimización	58
3.2. Optimización multiobjetivo	60
3.3. Extensión de la media proximal	67
Bibliografía	71

Introducción

"Dado que la forma del Universo entero es la más perfecta y, de hecho, la más sabiamente creada, absolutamente nada en el mundo ocurrirá sin que una minimización o maximización esté actuando". Leonard Euler 1744 .

En muchas situaciones reales se trata de utilizar determinados recursos en una cantidad limitada pero de la mejor manera, es decir que su uso cause el mayor provecho. La programación lineal estudia la optimización de una función lineal que satisface un conjunto de restricciones lineales de igualdad o desigualdad. La programación lineal es un modelo matemático que fue planteado por primera vez por George B. Dantzing en 1947 cuando era consejero matemático de la fuerza aérea de los Estados Unidos. Sabemos además que en 1939 Leonid V. Kantorovich ya había planteado y resuelto problemas de este tipo.

En aplicaciones de la optimización a la economía, teoría de control, problemas inversos etc, surgen problemas donde la función objetivo no siempre es diferenciable o casos en los cuales el problema no está bien puesto. Para resolver problemas como estos se utilizan técnicas en el contexto del análisis convexo, como los métodos de regularización para funciones convexas así como los métodos de punto proximal y lagrangeano aumentado entre otros.

Recientemente, en el año 2009, los profesores Bauschke, Lucet y Triens propusieron la media proximal, una novedosa técnica que tiene la propiedad de ser autodual respecto a la conjugada de Fenchel, que puede trabajar incluso con funciones de dominio disjunto, veremos que esta técnica puede ser aprovechada para manipular la envoltura de Goebel y probar su autodualidad respecto a la conjugada de Fenchel, además de tratar la optimización de varias funciones objetivo en el caso convexo o inclusive en el caso de ciertas funciones no necesariamente convexas aun cuando los dominios de estas funciones sean disjuntos.

Notaciones

Presentamos las notaciones utilizadas en el desarrollo de la teoría que vamos a exponer.

\mathbb{R}	el conjunto de los números reales.
$\overline{\mathbb{R}}$	es el conjunto $\mathbb{R} \cup \{-\infty, +\infty\}$
\mathbb{R}^n	el conjunto $\{x = (x_1, x_2, \dots, x_n) / x_i \in \mathbb{R}, \forall i = 1, 2, \dots, n\}$
\mathbb{R}_+^n	es el conjunto $\{x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \setminus x_i \geq 0, i = 1, 2, \dots, n\}$
$\langle \cdot, \cdot \rangle$	es el producto interno usual.
$\text{int}(\mathbb{X})$	denota el interior del conjunto $\mathbb{X} \subset \mathbb{R}^n$
i_S	función indicadora de un conjunto S .
$\text{dom} f$	dominio de una función f .
∂f	subdiferencial de una función f .
$\text{epi} f$	epígrafo de una función f .
$\Gamma_0(\mathbb{X})$	conjunto de las funciones convexas propias y semicontinuas inferiormente de $\mathbb{X} \subset \mathbb{R}^n$ en $\overline{\mathbb{R}}$.
Cl	clausura.
f^*	conjugada de Fenchel.
\mathcal{Q}	denota la función $\mathcal{Q}(x) = \frac{1}{2}\ x\ ^2$.
$f \sharp g$	convolución infimal o epi multiplicación de f y g .
$\mu \star f$	epi-multiplicación de μ no negativo con la función f
$\mathcal{P}_\mu(f, \lambda)$	media proximal de una función f con peso λ y parámetro μ donde $\mu \in \mathbb{R}$ $f = (f_1, \dots, f_n)$, $f_i \in \Gamma_0(\mathbb{X})$, $\lambda = (\lambda_1, \dots, \lambda_n)$, $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$, $1 \leq i \leq n$
$e_\lambda f$	envoltura de Moreau de una función f
$G_\lambda f$	envoltura de Goebel de una función f
$h_\lambda f$	λ cápsula proximal de una función f

Capítulo 1

Análisis convexo

Denotaremos por \mathbb{X} un subconjunto de \mathbb{R}^n , $\langle \cdot, \cdot \rangle$ el producto interno usual, norma $\| \cdot \|$ y $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ con estas notaciones revisaremos aspectos fundamentales del análisis convexo.

1.1. Conjuntos convexos

Definición 1.1. Sean $x_1, x_2 \in \mathbb{X}$, la *recta* que pasa por x_1, x_2 es el conjunto

$$(1.1) \quad \mathcal{L}_{x_1, x_2} = \{x/x = tx_1 + (1-t)x_2 ; t \in \mathbb{R}\}$$

y denotamos al *segmento de recta cerrado* de x_1 a x_2 como

$$[x_1, x_2] = \{x/x = tx_1 + (1-t)x_2 ; 0 \leq t \leq 1\}$$

Definición 1.2. Un conjunto $S \subseteq \mathbb{X}$ es *afín* si para cualesquiera $x_1, x_2 \in S$, se tiene $\mathcal{L}_{x_1, x_2} \subseteq S$

Definición 1.3. Un conjunto $S \subseteq \mathbb{X}$ es *convexo* si para cualesquiera $x_1, x_2 \in S$, se tiene $[x_1, x_2] \subseteq S$

Definición 1.4. Sean $a \in \mathbb{R}^n, a \neq 0$, $\alpha \in \mathbb{R}$ definimos el *hiperplano* como el conjunto:

$$H_a^\alpha = \{x \in \mathbb{R}^n / \langle x, a \rangle = \alpha\}$$

Se define:

Semiespacio positivo

$$H_a^{\alpha+} = \{x \in \mathbb{R}^n / \langle x, a \rangle > \alpha\}$$

Semiespacio negativo

$$H_a^{\alpha-} = \{x \in \mathbb{R}^n / \langle x, a \rangle < \alpha\}$$

Lema 1.1. Sea $\{\mathcal{C}_\beta\}_{\beta \in \mathcal{B}}$ una colección de conjuntos convexos tal que definimos $\mathcal{C} = \bigcap_{\beta \in \mathcal{B}} \mathcal{C}_\beta$, entonces \mathcal{C} es convexo.

Demostración. Sean $x_1, x_2 \in \mathcal{C}$ entonces $x_1, x_2 \in \mathcal{C}_\beta$ para todo $\beta \in \mathcal{B}$ luego $[x_1, x_2] \subseteq \mathcal{C}_\beta$ para todo $\beta \in \mathcal{B}$ de donde $[x_1, x_2] \subseteq \bigcap_{\beta \in \mathcal{B}} \mathcal{C}_\beta$, por lo tanto \mathcal{C} es convexo. \square

Ejemplo 1.1. Los siguientes conjuntos son convexos

1. Los conjuntos afines .

Sea A un conjunto afín, tomamos $x_1, x_2 \in A$ entonces $\mathcal{L}_{x_1, x_2} \subseteq A$, en particular $[x_1, x_2] \subseteq A$, por lo que el conjunto afín A es convexo.

2. Los semi espacios $H_a^{\alpha+}$ y $H_a^{\alpha-}$.

Probaremos que el semi espacio $H_a^{\alpha+}$ es convexo.

Tomamos $x_1, x_2 \in H_a^{\alpha+}$, entonces se tiene que $\langle x_1, a \rangle > \alpha$ y $\langle x_2, a \rangle > \alpha$ entonces para $\lambda \in [0, 1]$ se tiene que $\lambda x_1 + (1 - \lambda)x_2 \in H_a^{\alpha+}$ pues

$$\begin{aligned} \langle \lambda x_1 + (1 - \lambda)x_2, a \rangle &= \langle \lambda x_1, a \rangle + \langle (1 - \lambda)x_2, a \rangle \\ &= \lambda \langle x_1, a \rangle + (1 - \lambda) \langle x_2, a \rangle \\ &> \lambda \alpha + (1 - \lambda) \alpha \\ &> \alpha \end{aligned}$$

Así hemos probado que el semi espacio $H_a^{\alpha+}$ es convexo y en forma similar se tiene que $H_a^{\alpha-}$ es convexo.

3. El sistema de inecuaciones

$$\begin{cases} a_{1,1}x_1 & \cdots & +a_{1,n}x_n & \leq & b_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m,1}x_1 & \cdots & +a_{m,n}x_n & \leq & b_m \end{cases}$$

donde $a_{i,j}, b_j \in \mathbb{R}$ para $i \in \{1 \dots m\}$, $j \in \{1 \dots n\}$, se denota

$$Ax \leq b, \text{ donde } A = (a_{i,j}) \text{ y } b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Sea $S = \{x \in \mathbb{R}^n / Ax \leq b\}$ el conjunto solución del sistema de inecuaciones, S es la intersección de n semi espacios de la forma

$$S_i = \{x \in \mathbb{R}^n / A_i x \leq b_i\}$$

donde A_i denota la i -ésima fila de la matriz A , los cuales según el ejemplo anterior son convexos y por el Lema 1.1 el conjunto S es convexo.

Definición 1.5. El *interior* de un conjunto $C \subseteq \mathbb{R}^n$ es el conjunto

$$\text{int } C = \{x \in C \mid \text{existe } \epsilon > 0, B(x, \epsilon) \subseteq C\}$$

Un conjunto A será llamado *abierto* cuando $A = \text{int } A$; un conjunto B será llamado *cerrado* cuando su complemento sea abierto

Definición 1.6. La *cerradura* de un conjunto $C \subseteq \mathbb{R}^n$ es el conjunto cerrado más pequeño que contiene a C y es denotado por $\text{cl } C$

1.2. Función convexa

Definición 1.7. Sea una función $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ el *dominio efectivo* de f es

$$\text{dom}(f) = \{x \in \mathbb{X} \mid f(x) < \infty\}$$

Definición 1.8. Dada una función $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ el *conjunto de nivel inferior* o *subnivel* de parámetro $\gamma \in \mathbb{R}$ es un subconjunto de \mathbb{X} definido y denotado como

$$\Gamma_\gamma(f) = \{x \in \mathbb{X} \mid f(x) \leq \gamma\}$$

Definición 1.9. Sea $S \subseteq \mathbb{R}^n$, la *función indicadora* de S es dada por

$$i_S(x) = \begin{cases} 0 & , x \in S \\ +\infty & , x \notin S. \end{cases}$$

se verifica

$$(f + i_S)(x) = f(x) + i_S(x)$$

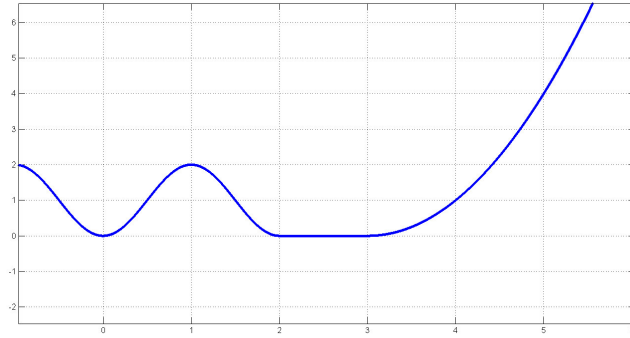
De aquí en adelante identificaremos a $(f + i_S)(x)$ con $f(x)$

$$\inf_{x \in S} f(x) = \inf_{x \in \mathbb{R}^n} \{f(x) + i_S(x)\}$$

Definición 1.10. Dada la función $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$, se define y se denota el *argumento mínimo* de la función f

$$\text{argmin } f = \begin{cases} \{x \in \mathbb{X} \mid f(x) = \inf_{\mathbb{X}} f\} & \text{si } \inf_{\mathbb{X}} f < +\infty \\ \phi & \text{si } \inf_{\mathbb{X}} f = +\infty. \end{cases}$$

$$\inf_{\mathbb{X}} f(x) = \inf_{\text{dom } f} f,$$

Figura 1.1: $\operatorname{argmin} f = \{0\} \cup [2, 3]$

Definición 1.11. Sea C un conjunto convexo no vacío en \mathbb{X} . Una función $f : C \rightarrow \overline{\mathbb{R}}$ es *convexa* si para cualesquiera $x_1, x_2 \in C$ y $0 \leq \alpha \leq 1$ se cumple:

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$$

La función f es llamada *estrictamente convexa* si la desigualdad es estricta para todo $x_1, x_2 \in C$, $x_1 \neq x_2$ y $0 < \alpha < 1$.

Teorema 1.2. (Desigualdad de Jensen)

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$. Entonces f es convexa si y solo si

$$f\left(\sum_{i=1}^l \alpha_i x_i\right) \leq \sum_{i=1}^l \alpha_i f(x_i)$$

donde $\alpha_i \geq 0$ y $\sum_{i=1}^l \alpha_i = 1$

Demostración. Vea [2], Teorema 1.1

□

Definición 1.12. Una función f definida en un conjunto convexo C es llamada *cóncava* si $-f$ es convexa.

Definición 1.13. Una función f definida en un conjunto convexo C es llamada *afín* si es finita, cóncava y convexa.

Ejemplo 1.2. Sea C un conjunto convexo y una función $f : C \rightarrow \mathbb{R} \cup \{+\infty\}$ definida por $f(x) = \langle a, x \rangle - b$ donde $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, se tiene que f es una función afín.

En efecto, sean $x \in C$ y $\theta \in \mathbb{R}$ entonces

$$\begin{aligned} f(\theta x + (1 - \theta)y) &= \langle a, \theta x + (1 - \theta)y \rangle - b \\ &= \theta(\langle a, x \rangle - b) + (1 - \theta)(\langle a, y \rangle - b) \\ &= \theta f(x) + (1 - \theta)f(y) \end{aligned}$$

Por lo tanto las condiciones de concavidad y convexidad son satisfechas y por lo tanto f es afín.

Ejemplo 1.3. Sean $f_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, ($i = 1, \dots, m$) funciones convexas y λ_i , escalares no negativos.

Probaremos que la función $g(x) = \sum_{i=1}^m \lambda_i f_i(x)$ es convexa.

Consideramos $x, y \in \mathbb{R}^n$, $\lambda \in \langle 0; 1 \rangle$; de la definición de g se tiene que

$$g(\lambda x + (1 - \lambda)y) = \sum_{i=1}^m \lambda_i f_i(\lambda x + (1 - \lambda)y)$$

como cada f_i es convexa, aplicando el Teorema 1.2, obtenemos que

$$f_i(\lambda x + (1 - \lambda)y) \leq \lambda f_i(x) + (1 - \lambda)f_i(y) \text{ para todo } i = 1, \dots, m$$

por lo tanto

$$\begin{aligned} g(\lambda x + (1 - \lambda)y) &= \sum_{i=1}^m \lambda_i f_i(\lambda x + (1 - \lambda)y) \\ &\leq \sum_{i=1}^m \{\lambda f_i(x) + (1 - \lambda)f_i(y)\} \\ &\leq \sum_{i=1}^m \lambda f_i(x) + (1 - \lambda) \sum_{i=1}^m f_i(y) \\ g(\lambda x + (1 - \lambda)y) &\leq \lambda g(x) + (1 - \lambda)g(y) \end{aligned}$$

Por lo tanto g es una función convexa por la definición de convexidad 1.11.

Definición 1.14. Sea $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ una función, el *epígrafo* de f es definido y denotado por

$$\text{epi}(f) = \{(x, y) / x \in \mathbb{X}, y \in \mathbb{R}, y \geq f(x)\}$$

Lema 1.3. Una función f definida en un convexo C es convexa si y solo si $\text{epi}(f)$ es convexo.

Demostración. Vea [2], Lema 1.1

□

Lema 1.4. Si f es una función convexa definida en un conjunto convexo C en \mathbb{R}^n entonces para cada γ el conjunto $\{x \in C / f(x) \leq \gamma\}$ es vacío o convexo.

Demostración. Vea [2], Lema 1.21 □

Definición 1.15. Sea una función $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$, el *hipógrafo* de f es definido y denotado por

$$\text{hipo}(f) = \{(x, y) / x \in \mathbb{X}, y \leq f(x)\}$$

Ejemplo 1.4. Considere las funciones $f_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ convexas para cada $i \in I$, donde I es un conjunto de índices arbitrario no vacío.

Probaremos que la función $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ definida por $g(x) = \sup_{i \in I} f_i(x)$ es convexa.

Se tiene $(x; \alpha) \in \text{epi}(g)$ si y solo si $g(x) \leq \alpha$, lo que equivale a $f_i(x) \leq \alpha$ para cada $i \in I$. Por lo tanto

$$(x, \alpha) \in \bigcap_{i \in I} \text{epi}(f_i)$$

tenemos $\text{epi}(g) = \bigcap_{i \in I} \text{epi}(f_i)$, además por hipótesis f_i es convexa para todo

$i \in I$ por el Lema 1.3 $\text{epi}(f_i)$ es convexo, además $\bigcap_{i \in I} \text{epi}(f_i)$ es convexa por

Lema 1.1, por lo tanto $\text{epi}(g)$ es convexo si y solo si g es convexa.

Ejemplo 1.5. Sea S un conjunto convexo, cerrado y no vacío en \mathbb{R}^n , probaremos que la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $f(y) = \inf_{x \in S} \|x - y\|$ es convexa.

Supongamos que f no es convexa, i.e., existen z_0, y_0 y λ_0 tales que

$$f(\lambda_0 z_0 + (1 - \lambda_0) y_0) > \lambda_0 f(z_0) + (1 - \lambda_0) f(y_0)$$

$$\inf_{x \in S} \|\lambda_0 z_0 + (1 - \lambda_0) y_0 - x\| > \lambda_0 \inf_{x \in S} \|x - z_0\| + (1 - \lambda_0) \inf_{x \in S} \|x - y_0\|$$

Sea $x_0 \in S$ tal que $\|x_0 - z_0\| = \inf_{x \in S} \|x - z_0\|$, tenemos que

$$\begin{aligned} \lambda_0 f(z_0) + (1 - \lambda_0) f(y_0) &= \inf_{x \in S} \lambda_0 \|x - z_0\| + \inf_{x \in S} (1 - \lambda_0) \|x - y_0\| \\ &\geq \inf_{x \in S} \{\|\lambda_0(z_0 - x_0)\| + \|(1 - \lambda_0)(y_0 - x)\|\} \\ &\geq \inf_{x \in S} \{\|\lambda_0(z_0 - x_0) + (1 - \lambda_0)(y_0 - x)\|\} \\ \lambda_0 f(z_0) + (1 - \lambda_0) f(y_0) &\geq \inf_{x \in S} \{\|\lambda_0 z_0 + (1 - \lambda_0) y_0 - [\lambda_0 x_0 + (1 - \lambda_0) x]\|\} \end{aligned}$$

Por lo tanto se obtiene que

$$\begin{aligned} f(\lambda_0 z_0 + (1 - \lambda_0) y_0) &> \lambda_0 f(z_0) + (1 - \lambda_0) f(y_0) \\ &\geq \inf_{x \in S} \{\|\lambda_0 z_0 + (1 - \lambda_0) y_0 - [\lambda_0 x_0 + (1 - \lambda_0) x]\|\} \end{aligned}$$

Para $\lambda_0 \neq 1$ tenemos que $\lambda_0 x_0 + (1 - \lambda_0)x \in S$, por ser convexo, definimos $L = \{l \in S / l = \lambda_0 x_0 + (1 - \lambda_0)x\}$, se observa que $L \subseteq S$, se tiene que

$$\begin{aligned} f(\lambda_0 z_0 + (1 - \lambda_0)y_0) &> \inf_{x \in S} \{\|\lambda_0 z_0 + (1 - \lambda_0)y_0 - [\lambda_0 x_0 + (1 - \lambda_0)x]\|\} \\ \inf_{x \in S} \|\lambda_0 z_0 + (1 - \lambda_0)y_0 - x\| &> \inf_{l \in L} \|\lambda_0 z_0 + (1 - \lambda_0)y_0 - l\| \end{aligned}$$

Es decir, se tiene que

$$\inf_{x \in S} \|\eta - x\| > \inf_{x \in L} \|\eta - x\|$$

Lo cual es una contradicción, por suponer que la función f no es convexa.

Definición 1.16. Una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ es llamada *función propia* si

1. f nunca toma el valor $-\infty$
2. Existe un $x \in \mathbb{R}^n$ tal que $f(x) < \infty$

Ejemplo 1.6. La función indicadora es una función propia.

Teorema 1.5.

Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, entonces se tiene que f es propia si y solo si epi f es un conjunto no vacío.

Demostración. Vea [1], Lema 1.21

□

Definición 1.17. Una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ es *semicontinua inferiormente* (s.c.i.) en $\bar{x} \in \mathbb{R}^n$ si

$$f(\bar{x}) \leq \liminf_{y \rightarrow \bar{x}} f(y)$$

Ejemplo 1.7. La función $f : [1, 2] \rightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} (x+1)^2 & , x \in \langle 1, 2] \\ 1 & , x = 1 \end{cases}$$

es semicontinua inferior en $[1, 2]$

Definición 1.18. La función $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es *coerciva* si

$$\lim_{\|x\| \rightarrow \infty} f(x) = +\infty$$

Ejemplo 1.8. La función $f(x_1, x_2) = x_1^4 + x_2^4 - 32x_2^2$ es coerciva.

Notación 1.1. El conjunto de las funciones convexas propias y semicontinuas inferiormente sobre \mathbb{X} será denotado por $\Gamma_0(\mathbb{X})$

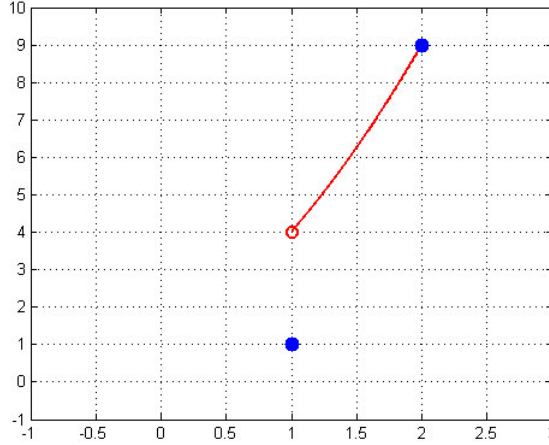


Figura 1.2: Gráfico de la función semicontinua del ejemplo 1.7

Definición 1.19. Una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ será llamada *función cerrada* si su epígrafo es un conjunto cerrado.

Definición 1.20. La *clausura inferior* de la función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ se denota $Cl(f)$ y se define

$$Cl(f) = \bar{f} = \sup\{g : \mathbb{X} \rightarrow \overline{\mathbb{R}} / g \text{ sci}, g \leq f\}$$

De la definición anterior se tiene que la clausura inferior o regularizada semicontinua inferior de la función f es el supremo de las funciones semicontinuas inferiormente mayoradas por la función f , es decir es la función cuyo epígrafo es la clausura del epígrafo de f en \mathbb{R}^{n+1} . La función clausura de f es llamada la cápsula semicontinua inferior de f y se denota por $Cl(f)$.

Teorema 1.6.

Sea una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, son equivalentes las siguientes afirmaciones

- i) El epígrafo es cerrado en \mathbb{R}^{n+1}
- ii) Para cada $\alpha \in \mathbb{R}$ el conjunto de nivel inferior $\{x / f(x) \leq \alpha\}$ es cerrado
- iii) f es semicontinua inferiormente en \mathbb{R}^n

Demostración. Haremos la demostración por partes, primeramente probaremos que i) implica ii), para lo cual tomamos una sucesión $x_n \rightarrow \hat{x}$ tal que $x_n \in \Gamma_\alpha f = \{x / f(x) \leq \alpha\}$, debemos probar que $\hat{x} \in \Gamma_\alpha f$.

En efecto como $x_n \rightarrow \hat{x}$ tal que $x_n \in \Gamma_\alpha f$ entonces $f(x_n) \leq \alpha$ por lo tanto $(x_n, \alpha) \in \text{epi}(f)$ es decir el $\text{epi}(f)$ es cerrado, así se tiene que $(\hat{x}, \alpha) \in \text{epi}(f)$ entonces $f(\hat{x}) \leq \alpha$ así $x_n \rightarrow \hat{x}$ tal que $x_n \in \Gamma_\alpha f$ entonces tenemos $\hat{x} \in \Gamma_\alpha f$

y así culminamos esta primera parte de la demostración .

En segundo lugar vamos a probar que *ii)* implica *iii)*, para lo cual vamos a considerar una sucesión $x_n \rightarrow x$ tal que $x_n \in \Gamma_\alpha f = \{x / f(x) \leq \alpha\}$ entonces $x \in \Gamma_\alpha f$, si suponemos que $\liminf_{n \rightarrow \infty} f(x_n) < f(x)$ es decir f no es s.c.i., entonces existe $\alpha \in \mathbb{R}$ tal que

$$(1.2) \quad \alpha < f(x)$$

pero como $f(x_n) \leq \alpha$ entonces se tiene $f(x) \leq \alpha$ lo cual es una contradicción con (1.2), por lo tanto $f(x) \leq \liminf_{n \rightarrow \infty} f(x_n)$, es decir se ha probado que la función f es semicontinua inferiormente en \mathbb{R}^n .

Por ultimo probaremos que *iii)* implica *i)*, para lo cual tomamos una sucesión $(x_n, t^n) \rightarrow (x, t)$ tal que $(x_n, t^n) \in \text{epi } f$ es decir que $f(x_n) \leq t^n$, ahora de la hipótesis *iii)* se tiene que f es sci., se tiene que $f(x) \leq \liminf_{n \rightarrow \infty} f(x_n)$ es decir $f(x) \leq t$, así $(x, t) \in \text{epi}(f)$ por lo tanto $\text{epi}(f)$ es cerrado, lo cual concluye la prueba. □

Teorema 1.7.

Sea la función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, se cumplen la siguientes afirmaciones

- i) $\text{epi}(Cl(f)) = Cl(\text{epi}(f))$. En particular, $Cl f$ es sci.
- ii) $Cl(\text{epi}(f))(x) = \liminf_{y \rightarrow x} f(y)$
- iii) f es sci. si y solo si $f \leq Cl(f)$ si y solo si $f = Cl(f)$

Demostración.

- i) Para probar la primera afirmación usamos el hecho que $\text{epi}(f) \subseteq \text{epi}(Cl(f))$ pues $Cl(f)$ es el menor cerrado que contiene a f ; puesto que $Cl(f)$ es sci., también se tiene que $\text{epi}(f) \subseteq Cl(\text{epi}(f))$ pues es el menor cerrado que contiene a $\text{epi}(f)$, por lo tanto se tiene que $\text{epi}(Cl(f)) = Cl(\text{epi}(f))$.
- ii) Vamos a probar la segunda proposición, para lo cual, recordemos que $Cl(f)$ es sci., entonces se tiene

$$(1.3) \quad Cl(f) \leq \liminf_{y \rightarrow x} (Cl(f)(y)) \leq \liminf_{y \rightarrow x} f(y)$$

puesto que $Cl(f)$ es el supremo de los minorantes de f ; definimos la función

$$h(x) = \liminf_{y \rightarrow x} f(y) = \sup \inf f(x)$$

por lo tanto

$$h \leq Cl(f)$$

es decir tenemos

$$(1.4) \quad \liminf_{y \rightarrow x} f(y) \leq Cl(f)$$

ahora tomando en cuenta las relaciones (1.3) y (1.4) tenemos probado nuestro resultado $Cl(f) = \liminf f$

iii) En esta parte de la demostración, probaremos que $Cl(f) = f$, para lo cual tenemos que

$$(1.5) \quad Cl(f) \leq f$$

como $Cl(f)$ es sci., entonces $f(x) \leq \liminf_{y \rightarrow x} f(y) = Cl(f)$ así tenemos

$$(1.6) \quad f \leq Cl(f)$$

por lo tanto de (1.5) y (1.6) tenemos $f(x) = Cl(f)$; ahora probaremos que f es sci., para lo cual si $f \leq Cl(f) = \liminf(f)$ entonces

$$f(x) \leq \liminf_{y \rightarrow x} f(y)$$

lo cual concluye la demostración.

□

1.2.1. Subgradiente

Definición 1.21. Sea una función $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{+\infty\}$. Decimos que f tiene un subgradiente en el punto $x_0 \in \text{dom}(f)$ si existe un vector $\zeta \in \mathbb{X}$ tal que se cumpla:

$$f(x) \geq f(x_0) + \langle \zeta, x - x_0 \rangle, \quad \forall x \in \mathbb{X}$$

El vector ζ es llamado *subgradiente* de f en x_0 .

Notación 1.2. Sea la función $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ denotaremos por $\partial f(x)$ el conjunto de todos los subgradientes de f en $x \in \text{dom}(f)$.

El conjunto

$$\partial f(x_0) = \{ \zeta \in \mathbb{X} / f(x) \geq f(x_0) + \langle \zeta, x - x_0 \rangle, \quad \forall x \in \mathbb{X} \}$$

es llamado el *subdiferencial* de f en el punto $x_0 \in \text{dom}(f)$

Teorema 1.8.

Sea f una función convexa definida en un conjunto convexo $C \subseteq \mathbb{R}^n$, entonces para cada punto $x_0 \in \text{int}(C)$ existe un vector $\zeta \in \mathbb{R}^n$ tal que:

$$f(x) \geq f(x_0) + \langle \zeta, x - x_0 \rangle, \quad \forall x \in C$$

Es decir $\partial f(x_0) \neq \emptyset$ en cada punto $x_0 \in \text{int}(C)$

Demostración. Vea [2], Teorema 2.1

□

Ejemplo 1.9. Sea $f(x) = |x|$, entonces

$$\partial f(x) = \begin{cases} \{-1\} & , x < 0 \\ \langle -1, 1 \rangle & , x = 0 \\ \{1\} & , x > 0 . \end{cases}$$

Ejemplo 1.10. Sea $C = \{(x_1, x_2) \in \mathbb{R} \times \mathbb{R} / 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$

$$f(x) = \begin{cases} 0 & , 0 \leq x_1 \leq 1, 0 < x_2 \leq 1 \\ \frac{1}{2} - |x_1 - \frac{1}{2}| & , 0 \leq x_1 \leq 1, x_2 = 0 \end{cases}$$

La función f tiene un único subgradiente $\xi = 0$ para todo $(x_1, x_2) \in \text{int}(C)$ pero la función f no es convexa en C .

Teorema 1.9.

Sea f una función definida en C un conjunto convexo en \mathbb{R}^n . Si en cada punto $x \in C$ existe el subgradiente de f , entonces la función f es convexa en C

Demostración. Vea [2], Teorema 2.2

□

1.2.2. Diferenciabilidad convexa

Teorema 1.10.

Sea f una función definida en un conjunto convexo $C \subseteq \mathbb{R}^n$ y sea $x \in \text{int}(C)$, entonces f es diferenciable en x si y solo si f tiene un único subgradiente en x . Además el único elemento en el subgradiente es $\nabla f(x)$.

Demostración. Vea [2], Teorema 3.1

□

Teorema 1.11.

Sea C un conjunto abierto y convexo de \mathbb{R}^n y sea f una función diferenciable en C . Entonces f es convexa si y solo si para cada $x_0 \in C$ se tiene que :

$$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle, \text{ para todo } x \in C$$

Además f es estrictamente convexa si y solo si para cada $x_0 \in C$ y $x \neq x_0$ en C se tiene:

$$f(x) > f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle, \text{ para todo } x \in C$$

Demostración. Vea [2], Teorema 3.2

□

Lema 1.12. Sea f de clase C^2 en un intervalo abierto $I \subseteq \mathbb{R}$. Entonces f es convexa en I si y solo si $f''(x) \geq 0$ para todo $x \in I$. Si $f''(x) > 0$ para toda $x \in I$ entonces f es estrictamente convexa.

Demostración. Vea [2], Lema 3.1

□

Ejemplo 1.11. La función $f(x) = \log\left(\frac{e^{ax}-1}{e^x-1}\right)$ con $a \geq 1$, $x \in \mathbb{R}$, es convexa.

Ejemplo 1.12. La función $f(x) = x \ln(x)$ donde $x > 0$, es convexa.

Ejemplo 1.13. La función $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $f(x) = x \ln\left(\frac{x}{p}\right) + p - x$ donde $x > 0$ y $p > 0$, es coerciva.

Definición 1.22. Una matriz $A \in \mathbb{R}^{n \times n}$ es *semidefinida positiva* si $x^t A x \geq 0$ para todo $x \in \mathbb{R}^n$, en este caso se denota $A \succeq 0$. Si $x^t A x > 0$ para todo $x \in \mathbb{R}^n$ la matriz A , es *definida positiva* en este caso se denota $A \succ 0$

Teorema 1.13.

Sea f de clase C^2 en un abierto y convexo D . Entonces f es convexa en D si y solo si la matriz hessiana

$$H(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

es semidefinida positiva para cada punto $x \in D$. Si $H(x)$ es definida positiva en cada x entonces f es estrictamente convexa.

Demostración. Vea [2], Teorema 3.3

□

Ejemplo 1.14. La función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por la siguiente regla de correspondencia $f(x_1, \dots, x_n) = \sum_{i=1}^n (2x_i - 1)(\ln x_i - \ln(1 - x_i))$; $x_i \in \langle 0, 1 \rangle$ es convexa.

En efecto para calcular la Hessiana de f es suficiente analizar la función

$$f(t) = (2t - 1)(\ln t - \ln(1 - t)), \quad t \in \langle 0, 1 \rangle$$

Calculamos la primera derivada tenemos

$$f'(t) = 2(\ln t - \ln(1 - t)) + (2t - 1) \left(\frac{1}{t} + \frac{1}{1 - t} \right)$$

Ahora calculamos la segunda derivada $f''(t) = \frac{1}{t^2(1-t)^2}$ luego

$$\nabla^2 f(x) = \begin{pmatrix} \frac{1}{x_1^2(1-x_1)^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{x_2^2(1-x_2)^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{x_n^2(1-x_n)^2} \end{pmatrix}$$

de donde se deduce $\nabla^2 f(x) \succ 0$ así tenemos que f es estrictamente convexa y en particular convexa.

Teorema 1.14. Composición de funciones afines

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $A \in \mathbb{R}^{n \times n}$, y $b \in \mathbb{R}^n$. Se define $g : \mathbb{R}^n \rightarrow \mathbb{R}$ por

$$g(x) = f(Ax + b)$$

donde $\text{dom } g = \{x / Ax + b \in \text{dom } f\}$, entonces si f es convexa, g es convexa.

Demostración. Vea [4], Teorema 3.2.2

□

1.2.3. Conjugada de Fenchel y inf-convolución

Definición 1.23. La *conjugada de Fenchel*, o conjugada convexa de la función $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ es definida como

$$f^*(y) = \sup_{x \in \text{dom}(f)} \{\langle x, y \rangle - f(x)\}$$

para todo $y \in \mathbb{R}^n$, es interesante observar que, dado que $\langle x, y \rangle - f(x)$ es afín con respecto a y para un x fijo, entonces f^* es el supremo de un conjunto de funciones convexas y por lo tanto es convexo, independientemente de si f es convexa o no.

Definición 1.24. Sea $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ se define $f^{**} = (f^*)^*$ como

$$f^{**}(x) = \sup_v \{\langle x, v \rangle - f^*(v)\}$$

El significado de la función conjugada f^* se puede dar en términos del epígrafo. Del gráfico se tiene que $(v, \beta) \in \text{epi}(f^*)$ si y solo si $f^*(v) \leq \beta$ lo cual equivale a $\langle v, x \rangle - f(x) \leq \beta$, para todo $(x, \alpha) \in \text{epi}(f)$ si y solo si $\langle v, x \rangle - \alpha \leq \beta$, para todo $(x, \alpha) \in \text{epi}(f)$, escribimos

$$\langle v, x \rangle - \beta \leq \alpha; \quad I_{v, \beta} = \langle v, x \rangle - \beta$$

así $(v, \beta) \in \text{epi}(f^*)$ si y solo si $I_{v, \beta} = \langle v, x \rangle - \beta \leq f$.

Entonces f^* describe una familia de funciones afines mayoradas por f , además nuestros cálculos revelan que $(v, \beta) \in \text{epi}(f^*)$ es decir $\beta \geq f^*(v)$ si y solo si

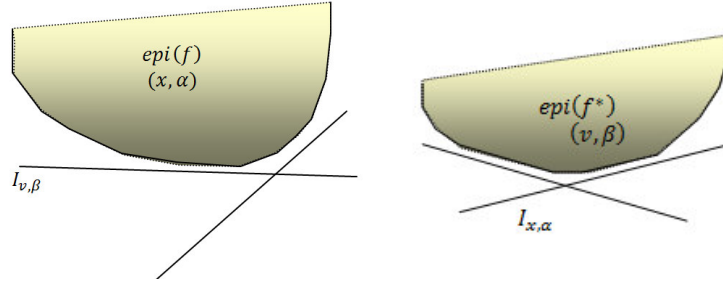


Figura 1.3: Función afín $l_{v,\beta}$ mayorada por f y función afín $l_{x,\alpha}$ es mayorada por $\text{epi}(f^*)$

$\beta \geq \langle x, v \rangle - \alpha$ de aquí $\beta \geq I_{x,\alpha}$ para todo $(x, \alpha) \in \text{epi } f$ en otras palabras f^* es punto a punto el supremo de todas las funciones afines $I_{x,\alpha} = \langle x, v \rangle - \alpha$ para $(x, \alpha) \in \text{epi } f$

Teorema 1.15.

Si $f_1 \leq f_2$, entonces se cumple $f_1^* \geq f_2^*$

Demostración. Sea $y \in \mathbb{X}$, como $f_1 \leq f_2$, entonces se cumple $\langle x, y \rangle - f_1(x) \geq \langle x, y \rangle - f_2(x)$ para todo $x \in \mathbb{X}$. Tomando el supremo se tiene

$$\sup_{x \in \mathbb{X}} \{ \langle x, y \rangle - f_1(x) \} \geq \sup_{x \in \mathbb{X}} \{ \langle x, y \rangle - f_2(x) \}$$

Por lo tanto se tiene que $f_1^* \geq f_2^*$. □

Teorema 1.16. Aproximación por funciones lineales afines

Una función $f : X \rightarrow \overline{\mathbb{R}}$ propia, convexa y cerrada es el supremo de la familia de funciones afines h tales que $h \leq f$.

Demostración. La prueba se encuentra en [3] Teorema 2.4.4. □

Teorema 1.17.

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, una función propia, se cumplen:

- i) Desigualdad de Young Fenchel $f(x) + f^*(p) \geq \langle p, x \rangle$, para cada $p \in \mathbb{R}^n$
- ii) $f^{**}(x) \leq f(x)$, para cada x , $f^{**} = f$ si y solo si f es convexa y cerrada.

Demostración.

- i) Para la demostración de la desigualdad de Young Fenchel, debemos recordar la definición de conjugada convexa $f^*(p) = \sup_x \{ \langle p, x \rangle - f(x) \}$ de aquí se tiene

$$(1.7) \quad f^*(p) \geq \langle p, x \rangle - f(x) \text{ si y solo si } f(x) + f^*(p) \geq \langle p, x \rangle$$

ii) Para esta demostración, consideramos la conjugada de la función f^* así tenemos $f^{**}(x) = \sup_y \{\langle y, x \rangle - f^*(y)\}$ pero de la ecuación (1.7) tenemos $\langle y, x \rangle - f^*(y) \leq f(x)$ de aquí se tiene

$$\sup_y \{\langle y, x \rangle - f^*(y)\} \leq f(x)$$

Por lo tanto

$$f^{**} \leq f$$

Ahora probaremos $f^{**} \geq f$, para ello, consideramos el conjunto Q formado por todas las funciones afines mayoradas por f , para cada $h \in Q$ tenemos $h(x) = \langle p, x \rangle - \alpha$ para cada x , por lo tanto

$$\alpha \geq \sup_x \{\langle p, x \rangle - f(x)\} = f^*(p)$$

$h(x) = \langle p, x \rangle - \alpha \leq f(x)$, para cada x , $\langle p, x \rangle - f(x) \leq \alpha$, para cada x

$$\sup_x \{\langle p, x \rangle - f(x)\} \leq \alpha$$

$f^*(p) \leq \alpha$ si y solo si $-\alpha \leq -f^*(p)$ y consecuentemente

$$\langle p, x \rangle - \alpha \leq \langle p, x \rangle - f^*(p)$$

así que $h(x) \leq \langle p, x \rangle - f^*(p)$, para cada x por lo tanto

$$(1.8) \quad \sup\{h / h \in Q\} \leq \sup_p \{\langle p, x \rangle - f^*(p)\} = f^{**}$$

Si f es convexa y cerrada, entonces es propia; por el Teorema 1.16, es justamente igual a la función del lado izquierdo de (1.8), es decir $f \leq f^{**}$ y como $f^{**} \leq f$ se tiene que $f^{**} = f$.

Recíprocamente si $f = f^{**}$ entonces f es la conjugada de f^* por lo tanto es convexa y cerrada.

□

Ejemplo 1.15. La conjugada de la función $f(x) = \exp(x)$ es

$$f^*(x) = \begin{cases} x \ln x - x & , x > 0 \\ 0 & , x = 0 \\ +\infty & , x < 0 \end{cases}$$

En efecto, pues la conjugada de f es $f^*(p) = \sup_{x \in \mathbb{R}} \{px - e^x\}$, ahora vamos a considerar los casos

1. Si $p > 0$, $f^*(p) = \sup_{x \in \mathbb{R}} \{px - e^x\}$

calculamos el máximo de $h(x) = px - e^x$

$$h(x) = px - e^x$$

$$h'(x) = p - e^x = 0 \leftrightarrow e^x = p \leftrightarrow x = \ln p$$

$$h(x) = p \ln p - p$$

Por lo tanto $f^* = p \ln p - p$ si $p > 0$

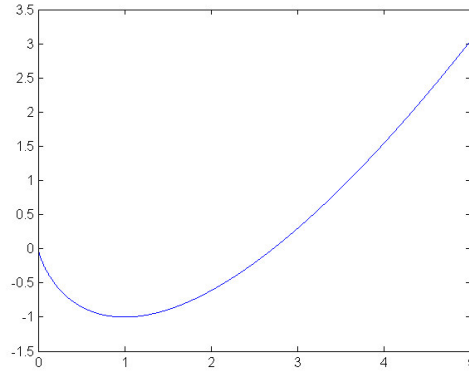


Figura 1.4: Gráfica de la función $f^*(p) = \sup_{x \in \mathbb{R}} \{px - e^x\}$, con $p > 0$

2. Si $p = 0$, $f^*(p) = \sup_{x \in \mathbb{R}} \{-e^x\} = 0$

3. Si $p < 0$, $f^*(p) = \sup_{x \in \mathbb{R}} \{px - e^x\} = +\infty$

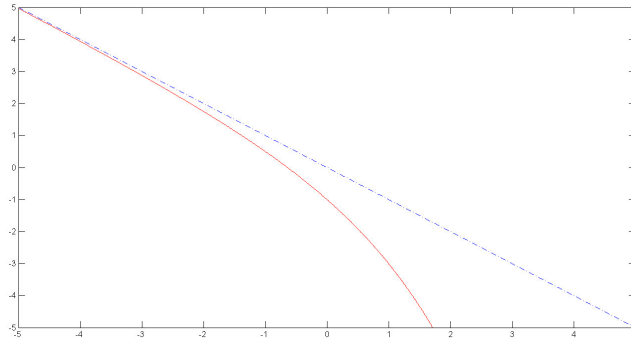


Figura 1.5: Gráfica de la función real de variable $x \in \mathbb{R}$, $px - e^x$, con $p < 0$

De 1, 2 y 3, tenemos la conjugada de la función f viene dada así

$$f^*(x) = \begin{cases} x \ln x - x & , x > 0 \\ 0 & , x = 0 \\ +\infty & , x < 0 . \end{cases}$$

Ejemplo 1.16. Sea $x \in \mathbb{X}$, la conjugada convexa de la función $q(x) = \frac{1}{2}\|x\|^2$ es

$$q^*(x) = \frac{1}{2}\|x\|^2 = q(x)$$

Además es la única función tal que $f^* = f$.

En efecto $f^*(p) = \sup_x \{\langle p, x \rangle - f(x)\} = \sup_x \left\{ \langle p, x \rangle - \frac{1}{2}\|x\|^2 \right\}$, definimos

$$h(x) = \langle p, x \rangle - \frac{1}{2}\|x\|^2$$

Si $\nabla h(x) = 0$ entonces $x = p$, este valor maximiza la función h , $f^*(p) = \frac{\|p\|^2}{2}$. Por otro lado, supongamos que f es una función convexa tal que $f^* = f$ entonces es una función propia y usando la desigualdad de Young - Fenchel se tiene $\langle x, x \rangle \leq f(x) + f^*(x) = 2f(x)$ es decir tenemos $q(x) \leq f(x)$. entonces por el Teorema 1.15 se tiene $f^* \leq q^* = q$. Como $f^* = f$ entonces tenemos $f \leq q$ y por lo tanto se tiene $q \leq f \leq q$, así concluimos $f = q$.

Notación 1.3. Denotaremos $\mathcal{Q}(x) = \frac{1}{2}\|x\|^2$ a la única función convexa tal que $\mathcal{Q}^* = \mathcal{Q}$

Ejemplo 1.17. Sea $f : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ una función definida por $f(x) = \frac{1}{p}|x|^p$
Para $p = 1$

$$f^*(x^*) = \begin{cases} 0 & \text{si } -1 \leq x^* \leq 1, \\ +\infty & \text{otro caso.} \end{cases}$$

y cuando $p > 1$

$$f^*(x^*) = \frac{1}{q}|x^*|^q \quad \text{donde } \frac{1}{p} + \frac{1}{q} = 1$$

En efecto.

Cuando $p = 1$, $f(x) = |x|$ y $f^*(x^*) = \sup_{x \in \mathbb{R}} \{xx^* - |x|\}$

$$\text{Si } x^* \geq 0, \quad f^*(x^*) = \sup_{x \in \mathbb{R}^+} \{xx^* - x\} = \sup_{x \in \mathbb{R}^+} \{x(x^* - 1)\} = \begin{cases} +\infty & \text{si } x^* > 1, \\ 0 & \text{si } x^* \leq 1. \end{cases}$$

$$\text{Si } x^* < 0, \quad f^*(x^*) = \sup_{x \in \mathbb{R}^-} \{xx^* + x\} = \sup_{x \in \mathbb{R}^-} \{x(x^* + 1)\} = \begin{cases} 0 & \text{si } x^* \geq -1, \\ +\infty & \text{si } x^* < -1. \end{cases}$$

Por lo tanto tenemos que:

$$f^*(x^*) = \begin{cases} 0 & \text{si } -1 \leq x^* \leq 1, \\ +\infty & \text{otros casos} \end{cases}$$

$$\text{Cuando } p > 1, f(x) = \frac{1}{p}|x|^p \text{ y } f^*(x^*) = \sup_{x \in \mathbb{R}} \left\{ xx^* - \frac{1}{p}|x|^p \right\}$$

Vamos a calcular el punto critico: $\frac{d}{dx} \left(xx^* - \frac{1}{p}|x|^p \right) = x^* - |x|^{p-1} \text{sgn}(x) = 0$
despejando $x^* = |x|^{p-1} \text{sgn}(x)$ de donde $|x^*| = |x|^{p-1}$ resulta $|x| = |x^*|^{\frac{1}{p-1}}$

Sustituyendo en la definición de la conjugada, tenemos

$$\begin{aligned}
 f^*(x^*) &= \sup_{x \in \mathbb{R}} \left\{ xx^* - \frac{1}{p}|x|^p \right\} \\
 &= x|x|^{p-1} \operatorname{sgn}(x) - \frac{1}{p}|x|^p \\
 &= |x|^p - \frac{1}{p}|x|^p \\
 &= \left(\frac{p-1}{p} \right) |x|^p \\
 &= \left(\frac{p-1}{p} \right) \left(|x^*|^{\frac{1}{p-1}} \right)^p \\
 f^*(x^*) &= \left(\frac{p-1}{p} \right) |x^*|^{\frac{p}{p-1}}
 \end{aligned}$$

Como $\frac{1}{p} + \frac{1}{q} = 1$ tenemos $q = \frac{p-1}{p}$, así concluimos que $f^*(x^*) = \frac{1}{q}|x^*|^q$

Ejemplo 1.18. Sea $f(x) = \frac{1}{p}|x - c|^p$ para alguna constante c con $p > 1$ entonces f es convexa.

En efecto, sea $g(x) = \left(\frac{1}{q}|x|^q \right)^* = \frac{1}{p}|x|^p$ con $\frac{1}{p} + \frac{1}{q} = 1$ entonces $g(x)$ es convexa pues es la conjugada de una función convexa, por lo tanto $f(x)$ es convexa.

Ejemplo 1.19. Si $f(x) = ax^2 + bx + c$ para $a > 0, c, b \in \mathbb{R}$ entonces se cumple $f^*(y) = \frac{1}{4a}(y - b)^2 - c$, y $f^{**} = f$.

En efecto, usando la definición de conjugada tenemos

$$\begin{aligned}
 f^*(p) &= \sup_x \{px - f(x)\}; a > 0 \\
 &= \sup_x \{px - ax^2 - bx - c\}; a > 0 \\
 f^*(p) &= \sup_x \{-ax^2 + (-b + p)x - c\}; a > 0
 \end{aligned}$$

Ahora definimos la función $h(x) = -ax^2 + (-b + p)x - c$, de aquí tenemos $h'(x) = -2ax - b + p = 0$ se obtiene $x = \frac{p-b}{2a}$ por lo tanto

$$\begin{aligned}
 f^*(p) &= -a \left(\frac{p-b}{2a} \right)^2 + (p-b) \left(\frac{p-b}{2a} \right) + c; a > 0 \\
 &= -a \frac{(p-b)^2}{4a^2} + \frac{(p-b)^2}{2a} - c; a > 0 \\
 &= (p-b)^2 \left(\frac{-1}{4a} - \frac{2}{4a} \right) - c; a > 0 \\
 &= \frac{(p-b)^2}{4a} - c; a > 0 \\
 f^*(y) &= \frac{(y-b)^2}{4a} - c; a > 0
 \end{aligned}$$

Ahora calcularemos f^{**}

$$\begin{aligned} f^{**}(p) &= \sup_x \{px - f^*(x)\} \\ &= \sup_x \left\{ px - \frac{(x-b)^2}{4a} + c \right\} \end{aligned}$$

Definimos la función $g(x) = px - \frac{(x-b)^2}{4a} + c$ derivando se tiene la función $g'(x) = p - \frac{(x-b)}{2a} = 0$ despejando se tiene $x = 2ap + b$ por lo tanto

$$\begin{aligned} f^{**}(p) &= p(2ap + b) - \frac{(2ap)^2}{4a} + c \\ &= ap^2 + bp + c, \text{ así se concluye que} \\ f^{**}(y) &= f(y) \end{aligned}$$

Teorema 1.18.

Sea $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ una función propia, $x \in \mathbb{X}$. Son equivalentes:

- i) $y \in \partial f(x)$
- ii) $f(x) + f^*(y) \leq \langle x, y \rangle$
- iii) $f(x) + f^*(y) = \langle x, y \rangle$.

Además $f \in \Gamma_0(\mathbb{X})$ se cumple la fórmula de reciprocidad de Legendre

$$y \in \partial f(x) \text{ si y solo si } x \in \partial f^*(y)$$

Demostración. vamos a realizar la demostración por partes, primeramente probaremos que i) implica ii), dado $z \in \mathbb{X}$ se tiene que si $y \in \partial f(x)$, se tiene por definición de subgradiente

$$\begin{aligned} f(z) &\geq f(x) + \langle z - x, y \rangle \\ &\geq f(x) + \langle z, y \rangle - \langle x, y \rangle \\ \langle x, y \rangle &\geq f(x) + \langle z, y \rangle - f(z) \\ &\geq f(x) + \sup_z \{ \langle z, y \rangle - f(z) \} \\ \langle x, y \rangle &\geq f(x) + f^*(y) \end{aligned}$$

En segundo lugar probaremos que ii) implica iii), para lo cual tomamos como hipótesis la desigualdad de Young Fenchel $f(x) + f^*(p) \geq \langle p, x \rangle$, para cada $x \in \mathbb{X}$ para cada $p \in \mathbb{X}^*$ y por la hipótesis ii) $f(x) + f^*(p) \leq \langle p, x \rangle$ por lo tanto se cumple $f(x) + f^*(p) = \langle p, x \rangle$ y así tenemos probada la proposición iii).

Por ultimo probaremos que iii) implica i), para lo cual usamos la hipótesis

$$f(x) + f^*(p) = \langle p, x \rangle$$

entonces usando la definición de conjugada para f se tiene

$$f(x) + \sup_z \{\langle p, z \rangle - f(z)\} = \langle x, p \rangle$$

aplicando la definición de supremo, se tiene la siguiente desigualdad

$$f(x) + \langle p, z \rangle - f(z) \leq \langle x, p \rangle$$

luego enviando todo los términos al primer miembro de la desigualdad

$$f(x) + \langle p, z - x \rangle - f(z) \leq 0$$

$$f(x) + \langle p, z - x \rangle \leq f(z), \quad \text{para todo } z$$

Luego $p \in \partial f(x)$. Además $x \in \partial f^*(y)$ si y solo si $f^*(y) + f^{**}(x) = \langle x, y \rangle$ además por hipótesis $f \in \Gamma_0(\mathbb{X})$ y de aquí se tiene que $x \in \partial f^*(y)$ si y solo si $f^*(y) + f(x) = \langle x, y \rangle$ es decir $x \in \partial f^*(y)$ si y solo si $y \in \partial f(x)$ \square

Teorema 1.19.

Sea $f : \mathbb{X} \rightarrow \mathbb{R}\{+\infty\}$ una función y las constantes $\alpha, \beta \in \mathbb{R}$. Se cumple que

i) si $\alpha > 0$ entonces $(\alpha f)^*(x^*) = \alpha f^*(\alpha^{-1}x^*)$ para cada $x^* \in \mathbb{X}$.

ii) si $\beta \neq 0$ entonces $(f(\beta \cdot))^*(x^*) = f^*(\beta^{-1}x^*)$ para cada $x^* \in \mathbb{X}$.

Demostración. Vea [10], Teorema 2.3.1. \square

Definición 1.25. Sea una función cóncava $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ definimos el dominio como $\text{dom}(g) = \{x \in \mathbb{R}^n / g(x) > -\infty\}$, decimos que g es una función propia si $\text{dom}(g) \neq \emptyset$ y $g(x) < +\infty$ para todo $x \in \mathbb{R}^n$, la conjugada de una función cóncava y propia g es definida como

$$g^*(s) = \inf_{x \in \mathbb{R}^n} \{\langle x, s \rangle - g(x)\}$$

La función g^* es cóncava y propia.

Teorema 1.20. Dualidad de Fenchel

Sea f una función propia y convexa en \mathbb{R}^n , y sea g una función cóncava y propia en \mathbb{R}^n . Se verifica que

$$\inf_{x \in \mathbb{R}^n} \{f(x) - g(x)\} = \sup_{y \in \mathbb{R}^n} \{g^*(y) - f^*(y)\}$$

si se cumple una de las condiciones:

i) $\text{ri}(\text{dom } f) \cap \text{ri}(\text{dom } g) \neq \emptyset$

ii) f y g son sci y $\text{ri}(\text{dom } f^*) \cap \text{ri}(\text{dom } g^*) \neq \emptyset$

Observación 1.1.

Este teorema transforma bajo ciertas condiciones, un problema de minimización en un problema de maximización.

Demostración. Vea [9], Teorema 31.1. □

Ejemplo 1.20. Dadas las funciones $f(x) = 0$ y $g(x) = -x^2$ de \mathbb{R} en $\overline{\mathbb{R}}$ vamos a calcular $f^*(p)$ y $g^*(p)$

$$(1.9) \quad f^*(p) = \sup_x \{\langle x, p \rangle - f(x)\} = \sup_x \{\langle x, p \rangle\} = i_{\{0\}}(p)$$

$$(1.10) \quad g^*(p) = \sup_x \{\langle x, p \rangle - g(x)\} = \sup_x \{\langle x, p \rangle + x^2\} = -\frac{p^2}{4}$$

También vamos a calcular $\inf_x \{f(x) - g(x)\}$ y $\sup_x \{g^*(x) - f^*(x)\}$, así tenemos

$$(1.11) \quad \inf_x \{f(x) - g(x)\} = \inf_x \{x^2\} = 0$$

$$(1.12) \quad \sup_x \{g^*(x) - f^*(x)\} = \sup_x \left\{-\frac{x^2}{4} - i_{\{0\}}(x)\right\} = 0$$

Además notamos que $ri(dom f) \cap ri(dom g) \neq \emptyset$ lo cual junto a las relaciones 1.11 y 1.12 verifican el teorema 1.20 es decir se tiene

$$\inf_{x \in \mathbb{R}^n} \{f(x) - g(x)\} = \sup_{y \in \mathbb{R}^n} \{g^*(y) - f^*(y)\}$$

Definición 1.26. Una función $K : C \times D \rightarrow \overline{\mathbb{R}}$, con $C \times D \subset \mathbb{R}^m \times \mathbb{R}^n$, será llamada *función silla de montar* (o cóncavo-convexo), siempre que:

1. Para cada $y \in D$, $K(\cdot, y)$ es cóncava, además
2. Para cada $x \in C$, $K(x, \cdot)$ es convexa .

Ejemplo 1.21. son funciones silla

1. $H(x, y) = \frac{1}{8}(-x^4 + y^4)$
2. $K(x, y) = -f(x) + g(y)$, donde f y g son funciones reales convexas

Definición 1.27. Para la función silla H el *dominio efectivo* es

$$dom H = dom_1 H \times dom_2 H$$

donde $dom_1 H = \{x \mid H(x, y) > -\infty, \forall y\}$, $dom_2 H = \{y \mid H(x, y) < +\infty, \forall x\}$
 Para la función silla H definimos la *conjugada inferior* y la *conjugada superior*

$$(1.13) \quad \underline{H}^*(p, q) = \sup_y \inf_x \{\langle p, x \rangle + \langle q, y \rangle - H(x, y)\}$$

$$(1.14) \quad \overline{H}^*(p, q) = \inf_x \sup_y \{\langle p, x \rangle + \langle q, y \rangle - H(x, y)\}$$

Definimos el *padre convexo* y el *padre concavo* de H

$$(1.15) \quad f(p, y) = \sup_x \{H(x, y) - \langle p, x \rangle\}, \quad g(x, q) = \inf_y \{H(x, y) - \langle q, y \rangle\}$$

Ejemplo 1.22. Dada una función $H(x, y) = xy$ calculamos

$$\begin{aligned} \underline{H}^*(p, q) &= \sup_y \inf_x \{\langle p, x \rangle + \langle q, y \rangle - xy\} = \sup_y \{\langle q, y \rangle + \inf_x \langle (p - y), x \rangle\} \\ &= \sup_y \{\langle q, y \rangle + i_{\{p\}}(y)\} = pq = \overline{H}^*(p, q) \end{aligned}$$

Definición 1.28. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ funciones propias y convexas, la convolución infimal o *inf-convolución* (o *epi-adición*), de f y g denotada $f \sharp g$ está definida como

$$(f \sharp g)(x) = \inf_u \{f(u) + g(x - u)\}$$

o

$$(f \sharp g)(x) = \inf_{x_1 + x_2 = x} \{f(x_1) + g(x_2)\}$$

además

$$(f_1 \sharp f_2 \sharp \cdots \sharp f_n)(x) = \inf_{x_1 + x_2 + \cdots + x_n = x} \{f(x_1) + \cdots + f(x_n)\}$$

.

Ejemplo 1.23. Dada una función f en \mathbb{R}^n y la función indicadora $i_{\{a\}}$ donde $a \in \mathbb{R}^n$ se tiene que

$$(f \sharp i_{\{a\}})(x) = \inf_p \{i_{\{a\}}(p) + f(x - p)\} = f(x - a)$$

es una traslación de la función f .

Ejemplo 1.24. Dada una función $f(x) = \|x\|$ norma Euclidiana y la función indicadora i_C donde C es un conjunto no vacío, calculamos la inf-convolución de ambas funciones

$$(f \sharp i_C)(x) = \inf_y \{\|x - y\| + i_C(y)\} = \inf_{y \in C} \{\|x - y\|\}$$

Ejemplo 1.25. Dadas las funciones $f(x) = x+2$ y $g(x) = x^2$, vamos a calcular la inf-convolución de ambas funciones

$$(1.16) \quad (f \sharp g)(p) = \inf_x \{f(x) + g(p-x)\} = \inf_x \{x+2 + (p-x)^2\}$$

Ahora consideramos la función $h(x) = (p-x)^2 + x+2$, para calcular su mínimo valor, vamos a derivar así obtenemos la función $h'(x) = 2x - 2p + 1 = 0$ de aquí se obtiene $x = \frac{2p-1}{2}$ por lo tanto reemplazando en (1.16) obtenemos la inf-convolución de las funciones $f(x) = x+2$ y $g(x) = x^2$

$$(f \sharp g)(p) = p + \frac{7}{4}$$

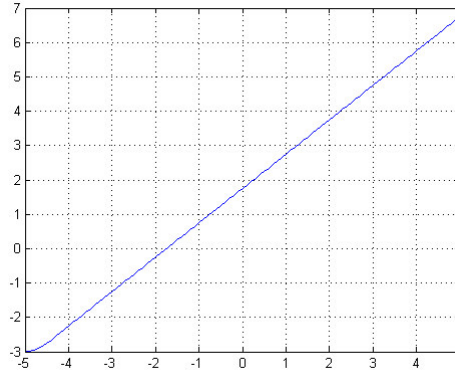


Figura 1.6: Gráfica de la función $f \sharp g$, donde $f(x) = x+2$ y $g(x) = x^2$

Verificaremos los cálculos realizados, para lo cual usaremos el siguiente código realizado con Matlab, el cual realizara el cálculo y gráfico de la función $f \sharp g$

```
f=@(x)(x+2); g=@(x)(x.^2);
x=-5:0.1:5;p=x;
esuma=zeros();j=1;
for i=p
esuma(j)=min(f(x)+g(i-x));
j=j+1;
end
plot(x,esuma)
grid on
```

Definición 1.29. Sea la función $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. Para un escalar no negativo μ , la *epi-multiplicación* $\mu \star f$ es definida por

$$(\mu \star f)(x) = \begin{cases} \mu f(\mu^{-1}x) & \text{si } \mu > 0 \\ i_{\{0\}}(x) & \text{si } \mu = 0 \end{cases}$$

Teorema 1.21.

Sea $f, f_1, \dots, f_n \in \Gamma_0(\mathbb{X})$. Entonces se cumplen:

- i) $\text{dom} (\alpha \star f) = \alpha(\text{dom } f)$
- ii) $\text{dom} (f_1 \sharp \dots \sharp f_n) = (\text{dom } f_1) + \dots + (\text{dom } f_n)$

Demostración.

- i) Sabemos que $\text{dom}(\alpha f) = \text{dom}(f)$ y también $\text{dom} f(\frac{x}{\alpha}) = \alpha \text{dom} f(x)$ de donde se tiene que $\text{dom} (\alpha \star f) = \alpha(\text{dom } f)$.
- ii) Para realizar esta segunda demostración vamos a proceder por doble inclusión, para lo cual tomamos $x \in \text{dom} (f_1 \sharp \dots \sharp f_n)$, entonces se tiene por la definición de episuma $x \in \text{dom} \left(\inf_{x_1+x_2+\dots+x_n=x} \{f(x_1) + \dots + f(x_n)\} \right)$ es decir $x = x_1 + x_2 + \dots + x_n$ con $x_i \in \text{dom}(f_i)$ para $i \in \{1, 2, \dots, n\}$; de aquí se tiene entonces que $x \in (\text{dom } f_1) + \dots + (\text{dom } f_n)$. Recíprocamente se tiene que si $x \in (\text{dom } f_1) + \dots + (\text{dom } f_n)$ entonces $x = x_1 + x_2 + \dots + x_n$ con $x_i \in \text{dom}(f_i)$ para $i \in \{1, 2, \dots, n\}$ esto quiere decir que $x \in \text{dom} \left(\inf_{x_1+x_2+\dots+x_n=x} \{f(x_1) + \dots + f(x_n)\} \right)$ por lo tanto se tiene $x \in \text{dom} (f_1 \sharp \dots \sharp f_n)$ y así concluye la demostración.

□

Teorema 1.22. Sea $\alpha \geq 0$, entonces se cumple que

- i) $(\alpha f)^* = \alpha \star f^*$
- ii) $(\alpha \star f)^* = \alpha f^*$
- iii) $(f_1 \sharp f_2 \sharp \dots \sharp f_n)^* = f_1^* + f_2^* + \dots + f_n^*$
- iv) Si $f \in \Gamma_0(\mathbb{X})$ se cumple que $f \sharp \mathcal{Q}$ y $f^* + \mathcal{Q}$ son conjugados uno del otro además se cumple $(f \sharp \mathcal{Q}) + (f^* \sharp \mathcal{Q}) = \mathcal{Q}$

Demostración.

- i) Vamos a demostrar el primer resultado, para esto vamos a considerar dos casos, para el primer caso consideramos $\alpha > 0$, de aquí se tiene

$$\begin{aligned}
 (\alpha f)^*(x^*) &= \sup_x \{ \langle x, x^* \rangle - \alpha f(x) \} \\
 &= \alpha \sup_x \left\{ \langle x, \frac{x^*}{\alpha} \rangle - f(x) \right\} \\
 &= \alpha f^* \left(\frac{x^*}{\alpha} \right)
 \end{aligned}$$

Para el segundo caso vamos a considerar $\alpha = 0$

$$(\alpha f)^* = \sup_x \{\langle x, x^* \rangle - 0\} = \begin{cases} 0 & \text{si } x^* = 0 \\ +\infty & \text{otro caso} \end{cases} = i_{\{0\}}(x^*) = (\alpha \star f^*)(x^*)$$

Por lo tanto tenemos $(\alpha f)^*(x) = (\alpha \star f^*)(x^*)$

ii) Para probar el segundo resultado vamos a usar la definición de epimultiplicación y de conjugada convexa

$$\begin{aligned} (\alpha \star f)^* &= \begin{cases} \sup_x \{\langle x, x^* \rangle - \alpha f(x/\alpha)\} & \text{si } \alpha > 0 \\ \sup_x \{\langle x, x^* \rangle - i_{\{0\}}(x)\} & \text{si } \alpha = 0 \end{cases} \\ &= \begin{cases} \alpha \sup_x \{\langle x/\alpha, x^* \rangle - f(x/\alpha)\} & \text{si } \alpha > 0 \\ 0 & \text{si } \alpha = 0 \end{cases} \\ (\alpha \star f)^* &= \alpha f^*(x^*) \end{aligned}$$

iii) Ahora vamos a mostrar la prueba del tercer resultado del teorema para lo cual vamos a usar la definición de epiadición y conjugada convexa y así se tiene

$$\begin{aligned} (f_1 \sharp f_2 \sharp \cdots \sharp f_n)^* &= \sup_x \left\{ \langle x, x^* \rangle - \inf_{x_1 + \cdots + x_n = x} \{f_1(x_1) + \cdots + f_n(x_n)\} \right\} \\ &= \sup_x \left\{ \langle x_1 + \cdots + x_n, x^* \rangle + \sup_{x_1 + \cdots + x_n = x} \{-f_1(x_1) - \cdots - f_n(x_n)\} \right\} \\ &= \sup_{x_1} \{\langle x_1, x^* \rangle - f_1(x_1)\} + \cdots + \sup_{x_n} \{\langle x_n, x^* \rangle - f_n(x_n)\} \\ (f_1 \sharp f_2 \sharp \cdots \sharp f_n)^* &= (f_1^* + \cdots + f_n^*)(x^*) \end{aligned}$$

iv) Por ultimo para la demostración de este resultado vamos a empezar mostrando que $f \sharp \mathcal{Q}$ y $f^* + \mathcal{Q}$ son conjugados uno del otro, usando la epiadición, tenemos $(f \sharp \mathcal{Q})^* = f^* + \mathcal{Q}$ tomando conjugada nuevamente se tiene $(f \sharp \mathcal{Q})^{**} = (f^* + \mathcal{Q})^*$ de aquí que se tiene el resultado $(f \sharp \mathcal{Q}) = (f^* + \mathcal{Q})^*$, así se ha demostrado que $f \sharp \mathcal{Q}$ y $f^* + \mathcal{Q}$ son conjugados uno del otro, ahora para demostrar que $(f \sharp \mathcal{Q}) + (f^* \sharp \mathcal{Q}) = \mathcal{Q}$ partimos de $(f \sharp \mathcal{Q})(w)$ y haciendo los cálculos tenemos

$$\begin{aligned} (f \sharp \mathcal{Q})(w) &= (f^* + \mathcal{Q})^*(w) \\ &= \sup_x \{\langle x, w \rangle - f^*(x) - \mathcal{Q}(x)\} \\ &= \sup_x \{\langle x, w \rangle - f^*(x) - \mathcal{Q}(x - w) + \mathcal{Q}(w) - \langle x, w \rangle\} \\ &= \sup_x \{\mathcal{Q}(w) - f^*(x) - \mathcal{Q}(x - w)\} \\ &= \mathcal{Q}(w) + \sup_x \{-f^*(x) - \mathcal{Q}(x - w)\} \\ &= \mathcal{Q}(w) - \inf_x \{f^*(x) + \mathcal{Q}(x - w)\} \\ (f \sharp \mathcal{Q})(w) &= \mathcal{Q}(w) - (f^* \sharp \mathcal{Q})(w) \end{aligned}$$

Es decir $(f \# \mathcal{Q}) + (f^* \# \mathcal{Q}) = \mathcal{Q}$, así queda probado el teorema. \square

Teorema 1.23.

Se cumplen las siguientes afirmaciones

- i) Si $\text{int}(\text{dom} f_1) \cap \dots \cap \text{int}(\text{dom} f_n) \neq \emptyset$, entonces $(f_1 + \dots + f_n)^* = f_1^* \# \dots \# f_n^*$
- ii) Si $\text{int}(\text{dom} f_1^*) \cap \dots \cap \text{int}(\text{dom} f_n^*) \neq \emptyset$, entonces $\text{epi}(f_1 \# \dots \# f_n) = \text{epi}(f_1) + \dots + \text{epi}(f_n)$
- iii) Para un escalar $\lambda > 0$ la epi-multiplicación, satisface $\text{epi}(\lambda \star f) = \lambda(\text{epi} f)$

Demostración. Vea [10], Teorema 2.8.7 \square

Ejemplo 1.26. Dada las funciones $f(x) = |x| + 1$ y $g(x) = (x - 1)^2$ vamos a mostrar las gráficas del $\text{epi}(f)$, $\text{epi}(g)$ y $\text{epi}(f \# g)$ para lo cual usaremos el código del ejemplo 1.25, considerando las nuevas funciones f y g

```
f=@(x)(abs(x)+1); g=@(x)((x-1).^2);
x=-5:0.1:7;p=x;y1=f(x);y2=g(x);
esuma=zeros();j=1;
for i=p
esuma(j)=min(f(x)+g(i-x));
j=j+1;
end
figure(1);fill(x,y1,'b');grid on ;axis([-5 5 0 6])
figure(2);fill(x,y2,'r');grid on;axis([-3 5 0 6])
figure(3);fill(x,esuma,'c');grid on ; axis([-5 7 0 6]);hold on
plot(x,y1,'b','LineWidth',2);plot(x,y2,'r','LineWidth',2)
```

Lema 1.24. *Se cumple*

$$(\lambda_1 \star (f_1 + \mathcal{Q}) \# \dots \# \lambda_n \star (f_n + \mathcal{Q}))^* = \lambda_1(f_1^* \# \mathcal{Q}) + \dots + \lambda_n(f_n^* \# \mathcal{Q})$$

Demostración. Para la demostración de este Lema, vamos a aplicar el Teorema 1.22 (ii) y (iii), Teorema 1.23 y el ejemplo 1.16, tenemos

$$\begin{aligned} (\lambda_1 \star (f_1 + \mathcal{Q}) \# \dots \# \lambda_n \star (f_n + \mathcal{Q}))^* &= (\lambda_1 \star (f_1 + \mathcal{Q}))^* + \dots + (\lambda_n \star (f_n + \mathcal{Q}))^* \\ &= \lambda_1(f_1 + \mathcal{Q})^* + \dots + \lambda_n(f_n + \mathcal{Q})^* \\ &= \lambda_1(f_1^* \# \mathcal{Q}^*) + \dots + \lambda_n(f_n^* \# \mathcal{Q}^*) \\ (\lambda_1 \star (f_1 + \mathcal{Q}) \# \dots \# \lambda_n \star (f_n + \mathcal{Q}))^* &= \lambda_1(f_1^* \# \mathcal{Q}) + \dots + \lambda_n(f_n^* \# \mathcal{Q}). \end{aligned}$$

\square

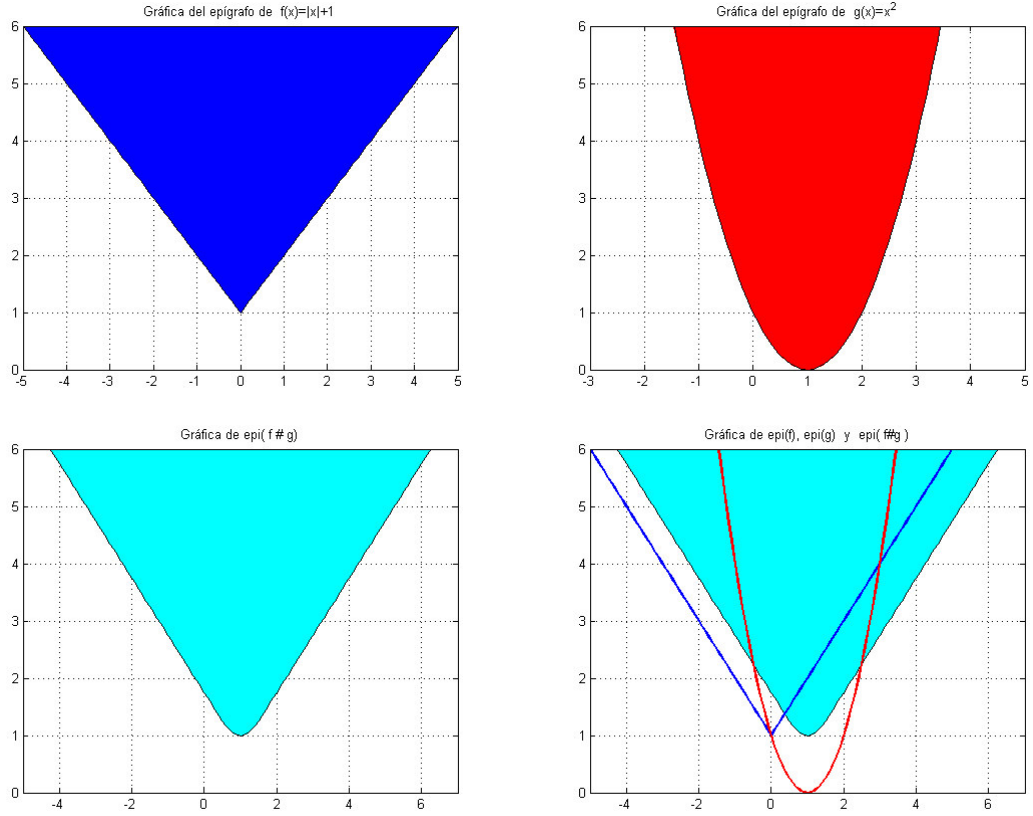


Figura 1.7: Gráfica del $\text{epi}(f)$, $\text{epi}(g)$ y $\text{epi}(f \# g)$ del ejemplo 1.26

Lema 1.25. Dadas las funciones $f_i : \mathbb{X} \rightarrow]-\infty, +\infty]$ para $1 \leq i \leq n$. Sean $f = (f_1, \dots, f_n)$, $x = (x_1, \dots, x_n)$, $\lambda_i \geq 0$ y $\tilde{f}(x) = \sum_{i=1}^n \lambda_i f_i(x_i)$, entonces se cumple $\tilde{f}^*(x^*) = \sum_{i=1}^n \lambda_i f_i^*(x_i^*)$

Demostración. Para la demostración, empezamos usando la definición de conjugada convexa

$$\begin{aligned}
 \tilde{f}^*(x^*) &= \sup_{x \in \mathbb{X}} \{ \langle x, x^* \rangle - \tilde{f}(x) \} \\
 &= \sup_{x=(x_1, \dots, x_n) \in \mathbb{X}} \left\{ \langle x_1, x_1^* \rangle + \dots + \langle x_n, x_n^* \rangle - \sum_{i=1}^n \lambda_i f_i^*(x_i^*) \right\} \\
 &= \lambda_1 \sup_{x_1} \{ \langle x_1, x_1^* / \lambda_1 \rangle - f_1(x_1) \} + \dots + \lambda_n \sup_{x_n} \{ \langle x_n, x_n^* / \lambda_n \rangle - f_n(x_n) \} \\
 &= \lambda_1 f_1^* \left(\frac{x_1^*}{\lambda_1} \right) + \dots + \lambda_n f_n^* \left(\frac{x_n^*}{\lambda_n} \right) \\
 \tilde{f}^*(x^*) &= \lambda_1 \star f_1^*(x_1^*) + \dots + \lambda_n \star f_n^*(x_n^*).
 \end{aligned}$$

□

Teorema 1.26.

Sea $g \in \Gamma_0(\mathbb{X})$, $h \in \Gamma_0(\mathbb{X})$ y para todo $x^* \in \mathbb{X}$ cumple

$$(g - h)^*(x^*) = \sup_{y^* \in \mathbb{X}} (g^*(y^*) - h^*(y^* - x^*))$$

Demostración. Vea [7], Teorema 2.2. □

Teorema 1.27.

Sea $g \in \Gamma_0(\mathbb{X})$, $\mu \geq 0$ se cumple que para todo $x^* \in \mathbb{X}$

$$(g - \mu \star \mathcal{Q})^*(x^*) = \mu(\mathcal{Q} - \mu^{-1}g^*)^* - \mu^{-1} \star \mathcal{Q}$$

Demostración. Usando los Teoremas 1.22 y 1.26 de ahí tenemos

$$\begin{aligned} (g - \mu \star \mathcal{Q})^*(x^*) &= \sup_{y^* \in \mathbb{X}} (g^*(y^*) - \mu \mathcal{Q}(y^* - x^*)) \\ &= \sup_{y^* \in \mathbb{X}} (g^*(y^*) - \mu \mathcal{Q}(y^*) - \mu \mathcal{Q}(x^*) + \mu \langle y^*, x^* \rangle) \\ &= -\mu \mathcal{Q}(x^*) + \sup_{y^* \in \mathbb{X}} (\mu \langle y^*, x^* \rangle - (\mu \mathcal{Q}(y^*) - \mathcal{Q}^*(y^*))) \\ &= -\mu \mathcal{Q}(x^*) + \mu \sup_{y^* \in \mathbb{X}} (\langle y^*, x^* \rangle - (\mathcal{Q}(y^*) - \mu^{-1} \mathcal{Q}^*(y^*))) \\ (g - \mu \star \mathcal{Q})^*(x^*) &= -(\mu^{-1} \star \mathcal{Q})(x^*) + \mu(\mathcal{Q} - \mu^{-1}g^*)^*(x^*). \end{aligned}$$

□

Operador de proximidad

Definición 1.30. El *operador de proximidad* o *el mapeo proximal* de una función $f \in \Gamma_0(\mathbb{X})$ es definido por

$$(1.17) \quad \text{Prox}_f x = \arg \min_{y \in \mathbb{X}} (f(y) + \mathcal{Q}(x - y)) \text{ para todo } x \in \mathbb{X}.$$

Teorema 1.28.

Para cualesquiera $x, p \in \mathbb{X}$ se tiene

$$p = \text{Prox}_f x \text{ si y solo si } x - p \in \partial f(p)$$

Demostración. Vea [9], Teorema 10.1 y el Ejercicio 10.2 □

Capítulo 2

La media proximal

Dadas dos funciones f_0 y f_1 , podemos construir una función $P(x, \lambda)$, de tal manera que $P(x, 0) = f_0(x)$ y $P(x, 1) = f_1(x)$ el método común suele ser el promedio ponderado es decir $P(x, \lambda) = \lambda_1 f_0(x) + \lambda_2 f_1(x)$ tal que λ_1, λ_2 son números reales no negativos de tal manera que $\lambda_1 + \lambda_2 = 1$, una aplicación que genera esta necesidad sería el problema de la optimización multiobjetivo, es decir calcular el $\min_x \{f_0(x), f_1(x)\}$, donde los objetivos se contraponen, y es importante considerar la prioridad que puede tener una función respecto a otra. Para tratar este tipo de problemas, vamos a considerar las siguientes funciones f_1, f_2, \dots, f_n se sabe que hay distintas metodológicas, que buscan una solución apropiada, y nuestro punto de inicio será el siguiente promedio ponderado de funciones denotado y definido de la siguiente manera

$$(2.1) \quad P(x, \lambda_1, \lambda_2, \dots, \lambda_n) = \lambda_1 f_1(x) + \dots + \lambda_n f_n(x), \quad \sum_{i=1}^n \lambda_i = 1, \quad \lambda_i \geq 0$$

Donde la intersección de los dominios de las funciones debe ser no vacío, justamente este requisito genera una gran dificultad, pues los dominios de las funciones objetivo pueden ser disjuntos, lo cual no nos va a permitir el tratamiento de la función (2.1). Para tratar esta dificultad en el año 2009, Bauschke y Lucet propusieron un promedio ponderado, llamado el promedio proximal o media proximal, denotado por $\mathcal{P}(f, \lambda)$, el cual supera la dificultad de calcular un promedio ponderado con las funciones f_1, \dots, f_n así sus dominios sean disjuntos. En este capítulo vamos a considerar $f_1, \dots, f_n \in \Gamma_0(\mathbb{X})$, y $\lambda_1, \dots, \lambda_n$ como números reales no negativos tales que $\sum_{i=1}^n \lambda_i = 1$

Definición 2.1. Media proximal

Sean $f = (f_1, \dots, f_n)$ y $\lambda = (\lambda_1, \dots, \lambda_n)$. La media proximal de peso λ con parámetro μ de n funciones f_i , $1 \leq i \leq n$ se define como

$$(2.2) \quad \mathcal{P}_\mu(f, \lambda) = \lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n + \mu \star \mathcal{Q}) - \mu \star \mathcal{Q}$$

Sea $x \in \mathbb{X}$, se tiene por definición de la epi-adición

$$\mathcal{P}_\mu(f, \lambda)(x) = \inf_{\sum_{i=1}^n x_i = x} \left\{ \lambda_1 \left[f_1\left(\frac{x_1}{\lambda_1}\right) + \frac{1}{\mu} \mathcal{Q}\left(\frac{x_1}{\lambda_1}\right) \right] + \cdots + \lambda_n \left[f_n\left(\frac{x_n}{\lambda_n}\right) + \frac{1}{\mu} \mathcal{Q}\left(\frac{x_n}{\lambda_n}\right) \right] \right\} - \frac{1}{\mu} \mathcal{Q}(x)$$

$$\mathcal{P}_\mu(f, \lambda)(x) = \frac{1}{\mu} \left(\inf_{\sum_{i=1}^n x_i = x} \left\{ \lambda_1 (\mu f_1 + \mathcal{Q})\left(\frac{x_1}{\lambda_1}\right) + \cdots + \lambda_n (\mu f_n + \mathcal{Q})\left(\frac{x_n}{\lambda_n}\right) \right\} - \mathcal{Q}(x) \right)$$

esto es, para todo $x \in \mathbb{X}$,

$$(2.3) \quad \mathcal{P}_\mu(f, \lambda)(x) = \frac{1}{\mu} \left(\inf_{\sum_{i=1}^n x_i = x} \sum_{i=1}^n \left\{ \lambda_i \left[\mu f_i\left(\frac{x_i}{\lambda_i}\right) + \mathcal{Q}\left(\frac{x_i}{\lambda_i}\right) \right] \right\} - \mathcal{Q}(x) \right)$$

Observación 2.1.

Cuando en la media proximal $\mu = 1$ escribiremos $\mathcal{P}(f, \lambda)(x)$ en lugar de $\mathcal{P}_1(f, \lambda)(x)$

Ejemplo 2.1. Vamos a calcular y graficar la función promedio ponderado de las funciones $f(x) = x + 2$ y $g(x) = x^2$ dada como $P(x, \lambda) = (1 - \lambda)f(x) + \lambda g(x)$ con $\lambda \in [0; 1]$, para todo $x \in [-3, 3]$

$$\begin{aligned} P(x, \lambda) &= (1 - \lambda)f(x) + \lambda g(x) \\ &= (1 - \lambda)(x + 2) + \lambda x^2 \\ P(x, \lambda) &= \lambda x^2 + (1 - \lambda)x + 2(1 - \lambda) \end{aligned}$$

Ahora vamos a graficar la función $P(x, \lambda)$ para diferentes valores de λ , por ejemplo $\lambda = 0$, $\lambda = 0.5$, $\lambda = 0.75$ y $\lambda = 1$; luego graficaremos la función promedio ponderado $P(x, \lambda)$ para 100 valores diferentes de λ en el intervalo $[0; 1]$ lo cual nos permite observar cómo la función promedio ponderado coincide con la función $f(x) = x + 2$ cuando $\lambda = 0$ y sigue variando hasta que cuando $\lambda = 1$, la función promedio ponderado coincide con la función $g(x) = x^2$; mostraremos el código que calcula y grafica el promedio ponderado para λ en el intervalo $[0; 1]$ y x en el intervalo $[-3, 3]$

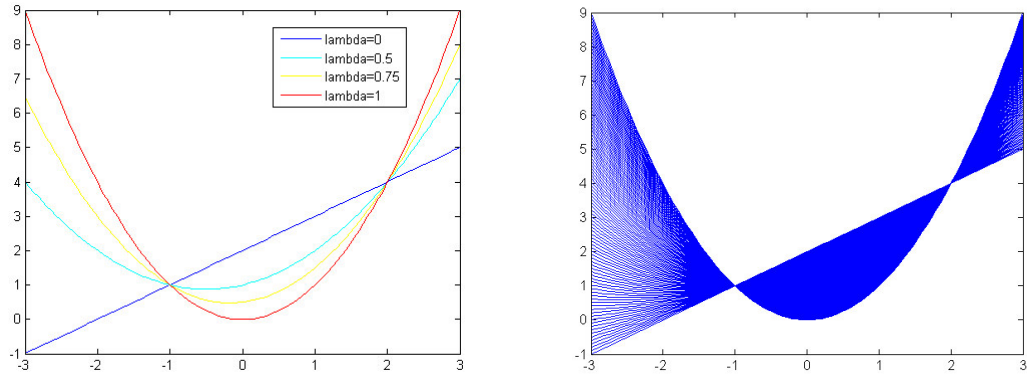


Figura 2.1: Función promedio ponderado $P(x, \lambda) = (1 - \lambda)f(x) + \lambda g(x)$

```
x=linspace(-3,3,100);
lamda=linspace(0,1,100);
y=x+2; plot(x,y)
hold on
for j=2:100 , y=(1-lamda(j)).*(x+2)+lamda(j).*x.^2; plot(x,y), end
```

Ejemplo 2.2. Calcularemos la media proximal de las funciones $f(x) = x + 2$ y $g(x) = x^2$, para lo cual usaremos la relación 2.3 con $\mu = 1$, $\lambda \in [0, 1]$ y $x \in [-3, 3]$

$$\begin{aligned} \mathcal{P}(f, g, \lambda)(p) &= \inf_x \left\{ \lambda \left[g\left(\frac{p-x}{\lambda}\right) + \mathcal{Q}\left(\frac{p-x}{\lambda}\right) \right] + (1-\lambda) \left[f\left(\frac{x}{1-\lambda}\right) + \mathcal{Q}\left(\frac{x}{1-\lambda}\right) \right] \right\} \\ &\quad - q(p) \\ &= \inf_x \left\{ \frac{3}{2} \lambda \left(\frac{x-p}{\lambda} \right)^2 + (1-\lambda) \left[\frac{x}{1-x} + 2 + \frac{1}{2} \left(\frac{x}{1-\lambda} \right)^2 \right] \right\} - \frac{p^2}{2} \end{aligned}$$

definimos la función $h(x) = \frac{3}{2} \lambda \left(\frac{x-p}{\lambda} \right)^2 + (1-\lambda) \left[\frac{x}{1-x} + 2 + \frac{1}{2} \left(\frac{x}{1-\lambda} \right)^2 \right]$ y procedemos a calcular su mínimo valor, para lo cual, calculamos la derivada $h'(x) = -(\lambda-1) \left(\frac{x}{(\lambda-1)^2} - \frac{1}{\lambda-1} \right) - \frac{3(2p-2x)}{2\lambda}$, luego resolviendo $h'(x) = 0$ despejamos $x = \frac{1-3p+3\lambda p-\lambda^2}{2\lambda-3}$, para reemplazar en la función $\mathcal{P}(f, g, \lambda)(p)$ y así se obtiene

$$\mathcal{P}(f, g, \lambda)(p) = -\frac{3(2p-1)^2}{8(2\lambda-3)} - \frac{p^2}{2} + \frac{3p}{2} - \frac{9\lambda}{4} + \frac{15}{8}$$

Ahora vamos a graficar la función $\mathcal{P}(f, g, \lambda)(p)$ para diferentes valores de λ por ejemplo si $\lambda = 0$, $\lambda = 0,5$, $\lambda = 0,75$ y $\lambda = 1$ se obtiene $\mathcal{P}(f, g, 0)(p) = p + 2$, $\mathcal{P}(f, g, 0,5)(p) = \frac{p^2}{4} + \frac{3p}{4} + \frac{15}{16}$, $\mathcal{P}(f, g, 0,75)(p) = \frac{p^2}{2} + \frac{p}{2} + \frac{7}{16}$ y $\mathcal{P}(f, g, 1)(p) = p^2$ respectivamente, se observa como cambia la función $\mathcal{P}(f, g, \lambda)(p)$, para observar mejor los cambios de la función $\mathcal{P}(f, g, \lambda)(p)$, presentaremos la gráfica de $\mathcal{P}(f, g, \lambda)(p)$ para 100 valores diferentes de λ en el intervalo $[0, 1]$, cuando p varía en el intervalo $[-3, 3]$, mostraremos el respectivo código, que nos calcula la función $\mathcal{P}(f, g, \lambda)(p)$ recibiendo como datos las funciones f y g

```

w=linspace(-3,3,100);
mprox=linspace(-3,3,100);
landa=linspace(0.0001,0.9999,100) ;
plot(0,0);hold on
for i=1:100
    for j=1:100
        landac(i)=1-landa(i);
        fun=@(x) (landa(i)*((x(1)/landa(i))+2 +(x(1)/landa(i))^2/2)+ ...
        landac(i)*((x(2)/landac(i))^2+ ((x(2)/landac(i))^2/2))-...
        ( w(j)^2/2)) ;
        x0 = [0 0 ];
        Aeq = [1 1];
        beq = w(j);
        [q,mprox(j)] = fmincon(fun,x0,[],[],Aeq,beq,[],[],[]);
    end
    plot(w,mprox);
end
end

```

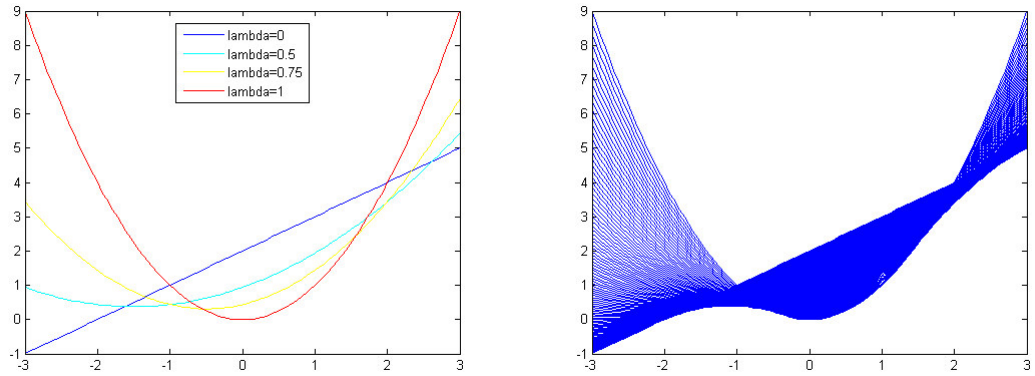


Figura 2.2: Gráfica de la función media proximal en el intervalo de $[-3, 3]$

2.1. Propiedades

Teorema 2.1.

La media proximal cumple las siguientes afirmaciones

i) $\mathcal{P}_\mu(f, \lambda)(x) = f_i(x)$, cuando $\lambda_i = 1$

ii) $\mathcal{P}_\mu(f, \lambda)(x) = \frac{1}{\mu} \left(\sum_{\lambda_i y_i = x} \inf \left\{ \sum_i \lambda_i \mu f_i(y_i) + \sum_i \lambda_i \mathcal{Q}(y_i) \right\} - \mathcal{Q}(x) \right)$

iii) $\mathcal{P}_\mu(f, \lambda) = (\lambda_1(f_1 + \mu^{-1}\mathcal{Q})^* + \dots + \lambda_n(f_n + \mu^{-1}\mathcal{Q})^*)^* - \mu^{-1}\mathcal{Q}$

Demostración.

- i) Para demostrar esta primera parte del Teorema recordemos la hipótesis $\lambda_k \geq 0$ para $k = 1, \dots, n$ y $\sum_{k=1}^n \lambda_k = 1$, basta tomar $\lambda_i = 1$ en la definición de media proximal, donde $i = 1, \dots, n$ y tenemos que $\mathcal{P}_\mu(f, \lambda)(x) = f_i(x)$, cuando $\lambda_i = 1$
- ii) Para realizar la demostración esta segunda parte del Teorema usamos el cambio de variables $y_i = \frac{x_i}{\lambda_i}$ en la relación (2.3) de la Definición 2.1 y tenemos inmediatamente concluida la prueba.
- iii) En esta tercera parte del Teorema, es necesario revisar el Teorema 1.21 (2) que nos afirma, para todo $i \in \mathbb{N}$, $\text{dom}(f_i^* \sharp \mathcal{Q}) = (\text{dom} f_i^*) + (\text{dom} \mathcal{Q}) = \mathbb{X}$, pues $\text{dom} \mathcal{Q} = \mathbb{X}$, entonces usando el Teorema 1.23 (1), el Teorema 1.22 (1) y el Teorema 1.17 se tiene

$$\begin{aligned}
\left(\sum_{i=1}^n \lambda_i (f_i + \mu^{-1} \mathcal{Q})^* \right)^* &= (\lambda_1 (f_1 + \mu^{-1} \mathcal{Q})^* + \dots + \lambda_n (f_n + \mu^{-1} \mathcal{Q})^*)^* \\
&= (\lambda_1 (f_1 + \mu^{-1} \mathcal{Q})^*)^* \sharp \dots \sharp (\lambda_n (f_n + \mu^{-1} \mathcal{Q})^*)^* \\
&= \lambda_1 \star (f_1 + \mu^{-1} \mathcal{Q})^{**} \sharp \dots \sharp \lambda_n \star (f_n + \mu^{-1} \mathcal{Q})^{**} \\
\left(\sum_{i=1}^n \lambda_i (f_i + \mu^{-1} \mathcal{Q})^* \right)^* &= \lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \sharp \dots \sharp \lambda_n \star (f_n + \mu \star \mathcal{Q}).
\end{aligned}$$

□

Teorema 2.2.

Se cumple la siguiente relación

$$\text{dom} \mathcal{P}_\mu(f, \lambda) = \lambda_1 \text{dom} f_1 + \dots + \lambda_n \text{dom} f_n$$

Demostración. Para la demostración de este resultado vamos a utilizar el Teorema 1.21 i) y la respectiva parte ii), se tiene

$$\begin{aligned}
\text{dom} \mathcal{P}_\mu(f, \lambda) &= \text{dom}(\lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \sharp \dots \sharp \lambda_n \star (f_n + \mu \star \mathcal{Q})) \\
&= \text{dom}(\lambda_1 \star (f_1 + \mu \star \mathcal{Q})) + \dots + \text{dom}(\lambda_n \star (f_n + \mu \star \mathcal{Q})) \\
&= \lambda_1 \text{dom}(f_1 + \mu^{-1} \mathcal{Q}(\frac{\cdot}{\mu})) + \dots + \lambda_n \text{dom}(f_n + \mu^{-1} \mathcal{Q}(\frac{\cdot}{\mu})) \\
\text{dom} \mathcal{P}_\mu(f, \lambda) &= \lambda_1 \text{dom}(f_1) + \dots + \lambda_n \text{dom}(f_n)
\end{aligned}$$

y así concluye la demostración.

□

2.1.1. Conjugada de Fenchel para la media proximal

Teorema 2.3.

Se cumple

$$(\mathcal{P}_\mu(f, \lambda))^* = \mathcal{P}_{\mu^{-1}}(f^*, \lambda)$$

Demostración. Para realizar esta demostración partiremos de la definición de media proximal y del Lema 1.24, tenemos

$$\begin{aligned} \mathcal{P}_\mu(f, \lambda) &= \lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n + \mu \star \mathcal{Q}) - \mu \star \mathcal{Q} \\ \mathcal{P}_\mu(f, \lambda) + \mu \star \mathcal{Q} &= \lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n + \mu \star \mathcal{Q}) \\ g = (\mathcal{P}_\mu(f, \lambda) + \mu \star \mathcal{Q})^* &= \lambda_1(f_1^* \sharp \mu \mathcal{Q}) + \cdots + \lambda_n(f_n^* \sharp \mu \mathcal{Q}) \end{aligned}$$

Ahora usando el Teorema 1.22 tenemos

$$\begin{aligned} \mathcal{Q} - \mu^{-1}g &= \lambda_1(\mathcal{Q} - \mu^{-1}(f_1^* \sharp \mu \mathcal{Q})) + \cdots + \lambda_n(\mathcal{Q} - \mu^{-1}(f_n^* \sharp \mu \mathcal{Q})) \\ &= \lambda_1(\mathcal{Q} - (\mu^{-1}f_1^* \sharp \mathcal{Q})) + \cdots + \lambda_n(\mathcal{Q} - (\mu^{-1}f_n^* \sharp \mathcal{Q})) \\ &= \lambda_1((\mu^{-1}f_1^*)^* \sharp \mathcal{Q}) + \cdots + \lambda_n((\mu^{-1}f_n^*)^* \sharp \mathcal{Q}) \\ &= \lambda_1(\mu^{-1} \star f_1 \sharp \mathcal{Q}) + \cdots + \lambda_n(\mu^{-1} \star f_n \sharp \mathcal{Q}) \end{aligned}$$

tomando conjugada y aplicando el Lema 1.24 tenemos

$$\begin{aligned} (\mathcal{Q} - \mu^{-1}g)^* &= (\lambda_1(\mu^{-1} \star f_1 \sharp \mathcal{Q}) + \cdots + \lambda_n(\mu^{-1} \star f_n \sharp \mathcal{Q}))^* \\ &= (\lambda_1(\mu^{-1} \star f_1 \sharp \mathcal{Q}))^* \sharp \cdots \sharp (\lambda_n(\mu^{-1} \star f_n \sharp \mathcal{Q}))^* \\ &= \lambda_1 \star (\mu^{-1} \star f_1 \sharp \mathcal{Q})^* \sharp \cdots \sharp \lambda_n \star (\mu^{-1} \star f_n \sharp \mathcal{Q})^* \\ &= \lambda_1 \star ((\mu^{-1} \star f_1)^* + \mathcal{Q}^*) \sharp \cdots \sharp \lambda_n \star ((\mu^{-1} \star f_n)^* + \mathcal{Q}^*) \\ &= \lambda_1 \star (\mu^{-1}f_1^* + \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (\mu^{-1}f_n^* + \mathcal{Q}) \end{aligned}$$

Ahora usando el Teorema 1.22 tenemos

$$\begin{aligned} \mu(\mathcal{Q} - \mu^{-1}g)^* &= \mu(\lambda_1 \star (\mu^{-1}f_1^* + \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (\mu^{-1}f_n^* + \mathcal{Q})) \\ &= \lambda_1 \star (f_1^* + \mu \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n^* + \mu \mathcal{Q}) \\ &= \lambda_1 \star (f_1^* + \mu^{-1} \star \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n^* + \mu^{-1} \star \mathcal{Q}) \end{aligned}$$

Combinando g con el Teorema 1.27 se tiene

$$\begin{aligned} (\mathcal{P}_\mu(f, \lambda))^* &= (\lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n + \mu \star \mathcal{Q}) - \mu \star \mathcal{Q})^* \\ &= \mu(\mathcal{Q} - \mu^{-1}((\lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n + \mu \star \mathcal{Q}))^*))^* - \mu^{-1} \star \mathcal{Q} \\ &= \lambda_1 \star (f_1^* + \mu^{-1} \star \mathcal{Q}) \sharp \cdots \sharp \lambda_n \star (f_n^* + \mu^{-1} \star \mathcal{Q}) - \mu^{-1} \star \mathcal{Q} \\ (\mathcal{P}_\mu(f, \lambda))^* &= \mathcal{P}_{\mu^{-1}}(f^*, \lambda) \end{aligned}$$

□

Teorema 2.4.

Para $\mu = 1$ se cumple

$$(\mathcal{P}(f, \lambda))^* = \mathcal{P}(f^*, \lambda)$$

Demostración. La demostración es inmediata si tenemos en cuenta la Observación 2.1 y el Teorema 2.3

□

Ejemplo 2.3. Sea $F = (f, f^*)$ y $\lambda = (\frac{1}{2}, \frac{1}{2})$, entonces $\mathcal{P}(F, \lambda) = \mathcal{Q}$.

En efecto, por el Teorema 2.4 tenemos que $(\mathcal{P}(F, \lambda))^* = \mathcal{P}(F^*, \lambda)$. Como $F^* = (f^*, f^{**}) = (f^*, f)$ entonces tenemos que $\mathcal{P}(F^*, \lambda) = \mathcal{P}(F, \lambda)$. Por lo tanto; usando el Ejemplo 1.16 tenemos que $(\mathcal{P}(F, \lambda))^* = \mathcal{Q}$.

Corolario 2.5.

$\mathcal{P}(f, \lambda)$ es convexa, semicontinua inferiormente y propia

Demostración. Por el Teorema 2.4 tenemos que

$$(\mathcal{P}(f, \lambda))^{**} = (\mathcal{P}(f^*, \lambda))^* = \mathcal{P}(f, \lambda)$$

Por lo tanto usando el Teorema 1.17 se tiene que $(\mathcal{P}(f, \lambda)) \in \Gamma_0(\mathbb{X})$

□

Definición 2.2. Para funciones propias, semicontinuas inferiormente en $\overline{\mathbb{R}}$ f y $\lambda > 0$, se tiene la *envoltura o sobre de Moreau* de la función f , que es denotado por $e_\lambda f$ y se define inicialmente como

$$ef(x) = \inf_u \left\{ f(u) + \frac{1}{2} \|x - u\|^2 \right\}$$

la cual luego es modificada por Yosida, agregando el parámetro λ convirtiéndose en la regularización Moreau-Yosida o simplemente Moreau, conocida por la expresión siguiente:

$$e_\lambda f(x) = \inf_u \left\{ f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right\}$$

Por lo tanto $e_\lambda f$ es la *inf*-convolución de la función f con la función

$$\mathcal{Q}(x) = \frac{\|x\|^2}{2}$$

entonces con nuestra notación tenemos $e_\lambda f(x) = f^\sharp(\lambda \star \mathcal{Q}(x))$.

El mapeo proximal $P_\lambda f(x)$ es un conjunto de mapeos dado por

$$P_\lambda f(x) = \arg \min_u \left\{ f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right\}$$

Observe que el método de punto proximal clásico dado como

$$P_{\lambda_k} f(x) = \arg \min_u \left\{ f(u) + \frac{1}{2\lambda_k} \|x - u\|^2 \right\}$$

sustituye el parámetro λ por una sucesión $\{\lambda_k\}$ acotada y de números positivos. Martinet, en 1970, introdujo este método en el estudio de la programación convexa, y posteriormente Rockafellar generalizó la técnica para operadores monótonos maximales.

Este método de punto proximal, debe ser visto primeramente como un método teórico, ya que en su aspecto práctico nos lleva a la resolución de una sucesión de subproblemas relativamente similares al original, que tendrán que ser resueltos con el apoyo de otros métodos; sin embargo, hay situaciones en las que es conveniente reemplazar un problema de minimización por una sucesión de subproblemas, como es el caso en el que se tiene un problema original con restricciones y se convierte en una sucesión de subproblemas irrestrictos, situación que se presenta en el método de Lagrangeano aumentado, usando la teoría de dualidad de Fenchel.

Definición 2.3. Una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ es de proximidad acotada si existe $\lambda > 0$ tal que $e_\lambda f(x) > -\infty$ para $x \in \mathbb{R}^n$.

Para una función f de proximidad acotada, se tiene que el supremo de todos los λ es llamado el *umbral* λ_f de proximidad acotada para f .

Teorema 2.6.

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ una función propia, sci y de proximidad acotada con umbral λ_f . Entonces, para $0 < \lambda \leq \mu < \lambda_f$, se cumplen las siguientes afirmaciones.

- i) $e_\mu f \leq e_\lambda f \leq f$. Por lo tanto, cuando $\lambda \rightarrow 0^+$, las envolturas $e_\lambda f$ forman una familia creciente de funciones acotadas por f y

$$\sup_{\lambda > 0} e_\lambda f(x) = f(x) \text{ para todo } x$$

ii) $e_\lambda(e_\mu f) = e_{\lambda+\mu} f$

iii) $\inf e_\lambda f = \inf f$ y $\operatorname{argmin} e_\lambda f = \operatorname{argmin} f$

Demostración.

- i) Para probar este primer resultado, recordemos que si $0 < \lambda \leq \mu$, entonces $f(u) + \frac{1}{2\mu} \|x - u\|^2 \leq f(u) + \frac{1}{2\lambda} \|x - u\|^2$, para todo u , por lo tanto tomando el ínfimo respecto de u se tiene $e_\mu f \leq e_\lambda f$. La otra desigualdad sigue de tomar $u = x$ en la definición de la envoltura de Moreau, tenemos $e_\lambda f \leq f$
- ii) Demostraremos el segundo resultado, para lo cual, adoptaremos la notación $\mathcal{Q}_\lambda(u) = \frac{1}{2\lambda} |u|^2$ tenemos que $e_\mu f(x) = f \# \mathcal{Q}_\lambda(x)$

$$\begin{aligned} e_\lambda(e_\mu f) &= e_\mu f \# \mathcal{Q}_\lambda(x) \\ &= f \# (\mathcal{Q}_\mu(x) \# \mathcal{Q}_\lambda(x)) \\ &= f \# (\mathcal{Q}_{\lambda+\mu}(x)) \\ &= e_{\lambda+\mu} f \end{aligned}$$

iii) Vamos a demostrar el ultimo resultado, para lo cual, deducimos de i) que $\inf e_\mu f \leq \inf f$. Por otro lado tenemos que $f(u) \leq f(u) + \frac{1}{2\lambda} \|x - u\|^2$, para cada x y cada u , por lo que tomando el ínfimo en ambos lados, tenemos $\inf f \leq e_\lambda f$, esto concluye la primera parte $\inf e_\lambda f = \inf f$, ahora para la segunda igualdad podemos proceder así

$$\operatorname{argmin} e_\lambda f = \partial(e_\mu f)^*(0) = \partial(f^* + \mu q)(0) = (\partial f^* + \mu Id)(0) = \partial f^*(0) = \operatorname{argmin} f$$

□

Ejemplo 2.4. Vamos a mostrar la gráfica de la envoltura de Moreau para la función

$$f(x) = ||x| - 1|$$

cuando $\lambda = 2$ y luego para $\lambda = 4$

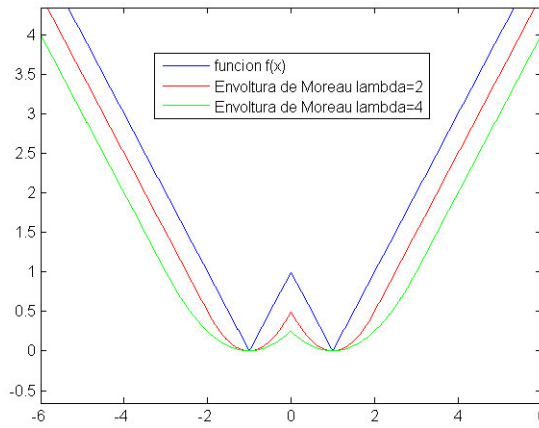


Figura 2.3: Función $f(x) = ||x| - 1|$ y su envoltura de Moreau, $\lambda = 2$, $\lambda = 4$

Ejemplo 2.5. Vamos a mostrar la gráfica de la envoltura de Moreau para la función $h(x)$

$$h(x) = \begin{cases} -2x + 10 & , 0 \leq x < 3, \\ 2x - 2 & , 3 \leq x < 4, \\ -4x + 22 & , 4 \leq x < 5, \\ 2 & , 5 \leq x < 7, \\ 8x - 54 & , 7 \leq x < 8, \end{cases}$$

cuando $\lambda = 2$

Teorema 2.7.

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ una función propia, sci, de proximidad acotada con umbral positivo t_f . Entonces para cada $\lambda \in \langle 0, 1 \rangle$ el conjunto P_λ es no vacío y compacto, donde el valor $e_\lambda f(x)$ es finito y depende continuamente en $\langle \lambda, x \rangle$, con

$$\lim_{\lambda \rightarrow 0^+} e_\lambda f(x) = \sup_{\lambda > 0} e_\lambda f(x) = f(x) \text{ para todo } x$$

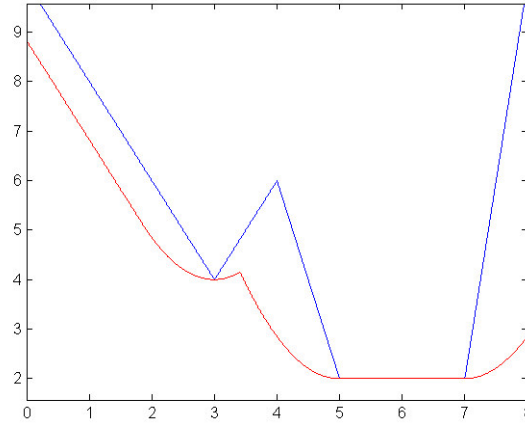


Figura 2.4: Función $h(x)$ y su envoltura de Moreau, $\lambda = 2$

Además si $x_m \rightarrow x$ y $\lambda_m \rightarrow 0^+$ en $\langle 0, t_f \rangle$ de tal manera que la sucesión $\{\frac{\|x_m - x\|}{\lambda_m}\}_{m \in \mathbb{N}}$ es acotada, entonces $e_{\lambda}f$ converge continuamente a f , es decir

$$e_{\lambda_m}f(x_m) \rightarrow f(x)$$

Demostración. Vea [10], teorema 1.25

□

Teorema 2.8. Cápsula proximal

Para $f : \mathbb{R} \rightarrow \overline{\mathbb{R}}$, $\lambda > 0$ la λ -cápsula proximal es la función $h_{\lambda}f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ definida por el supremo de las funciones cuadráticas de la forma $\alpha - \frac{1}{2\lambda}|x - w|^2$ que son mayoradas por f , esta $h_{\lambda}f$ satisface la desigualdad $e_{\lambda}f \leq h_{\lambda}f \leq f$ y se relaciona con la envoltura de Moreau $e_{\lambda}f$ por

$$h_{\lambda}f(x) = \sup_w \left\{ e_{\lambda}f(w) - \frac{1}{2\lambda}|x - w|^2 \right\}, \quad h_{\lambda}f = -e_{\lambda}(-e_{\lambda}f)$$

Demostración. Para realizar esta demostración consideramos la notación $\mathcal{Q}_{\lambda}(u) = \frac{1}{2\lambda}|u|^2$ para definir una nueva función $q_{\lambda,w,\alpha}(x) = \alpha - \mathcal{Q}_{\lambda}(x - w)$ donde $w \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ son parámetros. Por definición $h_{\lambda}f$ es el supremo de las funciones cuadráticas $q_{\lambda,w,\alpha}(x) = \alpha - \mathcal{Q}_{\lambda}(x - w)$ mayoradas por f . Estas funciones son continuas, así se tiene que $h_{\lambda}f$ es sci. Tenemos que :

$q_{\lambda,w,\alpha} \leq f$ es decir $\alpha - \mathcal{Q}_{\lambda}(x - w) \leq f$ sí y sólo sí $\alpha \leq f(x) + \mathcal{Q}_{\lambda}(x - w)$, para todo $x \in \mathbb{R}^n$, por la definición de la envoltura de Moreau tenemos que $\alpha \leq e_{\lambda}f(w)$, por lo tanto se tiene que $\alpha - \mathcal{Q}_{\lambda}(u) \leq e_{\lambda}f(w) - \mathcal{Q}_{\lambda}(u) \leq f(w)$ así se tiene que $h_{\lambda}f(x) = \sup_w \{ e_{\lambda}f(w) - \frac{1}{2\lambda}|x - w|^2 \}$ esto quiere decir que $h_{\lambda}f(x) = -\inf_w \{ -e_{\lambda}f(w) + \frac{1}{2\lambda}|x - w|^2 \}$ es decir se tiene que $h_{\lambda}f = -e_{\lambda}(-e_{\lambda}f)$. La desigualdad se obtiene de la relación de la capsula proximal y la envoltura de Moreau. □

Teorema 2.9.

Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ una función semicontinua inferiormente, propia y convexa, se cumplen las siguientes propiedades para cada $\lambda > 0$.

1. $P_\lambda f(x) \rightarrow P_{\bar{\lambda}} f(\bar{x})$ cuando $(\lambda, x) \rightarrow (\bar{\lambda}, \bar{x})$.
2. La función $e_\lambda f$ es convexa y continuamente diferenciable y su gradiente es $\nabla e_\lambda f = \frac{1}{\lambda}[x - P_\lambda f(x)]$

Demostración. Vea [10], Teorema 2.26

□

2.1.2. Aproximación regular

En esta sección se introduce una regularización que se puede aplicar a un amplio espectro de funciones. Su definición se basa en el envoltorio de Moreau. Una propiedad notable de esta regularización es su auto-dualidad con respecto a la conjugación de Fenchel, cuando se restringe a la clase de funciones propias, convexas y sci. Por otra parte, una extensión de esta regularización conserva las mismas características cuando se aplica a la clase de funciones silla de montar, propias y cerradas.

Definición 2.4. Sea $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ una función propia, sci y de proximidad acotada con umbral λ_f . Para cada $\lambda \in \langle 0, \lambda_f \rangle$, se define la *envoltura de Goebel o aproximación regular* $G_\lambda f$ como

$$(2.4) \quad G_\lambda f(x) = (1 - \lambda^2)e_\lambda f(x) + \frac{\lambda}{2}\|x\|^2$$

Teorema 2.10.

Cuando $\lambda \rightarrow 0^+$, se cumple: $G_\lambda f \rightarrow f$

Demostración. La demostración es directa a partir del Teorema 2.7

□

Observación 2.2. Es importante señalar, que G_λ no conserva otras características de la envoltura de Moreau. Ninguno de los resultados del Teorema 2.6 se cumple si se sustituye $e_\lambda f$ por $G_\lambda f$

Ejemplo 2.6. Sea la función $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x$, se tiene que $e_\lambda f(x) = x - \frac{\lambda}{2}$ y $G_\lambda f(x) = \frac{\lambda x^2}{2} + (1 - \lambda^2)x - \frac{\lambda(1 - \lambda^2)}{2}$. En efecto, sea $f(x) = x$ se tiene que $e_\lambda f(x) = \inf_u \{f(u) + \frac{1}{2\lambda}|x - u|^2\}$. Definimos la función

$$h(u) = f(u) + \frac{1}{2\lambda}|x - u|^2$$

derivando tenemos $h'(u) = 1 + \frac{u-x}{\lambda} = 0$ de aquí $u = x - \lambda$ por lo tanto

$$\begin{aligned} e_{\lambda}f(x) &= x - \lambda + \frac{(-\lambda)^2}{2\lambda} \\ &= x - \lambda + \frac{\lambda}{2} \\ e_{\lambda}f(x) &= x - \frac{\lambda}{2} \end{aligned}$$

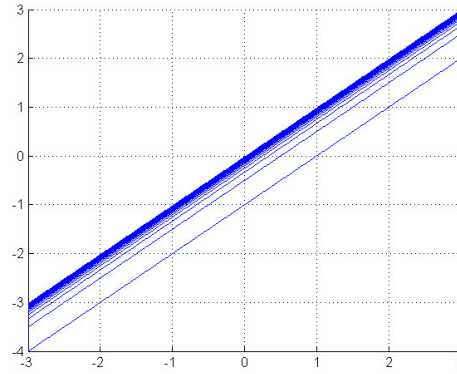


Figura 2.5: $e_{\lambda}f(x) = x - \frac{\lambda}{2}$

Por otro lado tenemos que

$$\begin{aligned} G_{\lambda}f(x) &= (1 - \lambda^2)e_{\lambda}f(x) + \frac{\lambda}{2}|x|^2 \\ &= (1 - \lambda^2)\left(x - \frac{\lambda}{2}\right) + \frac{\lambda}{2}x^2 \\ G_{\lambda}f(x) &= \frac{\lambda}{2}x^2 + (1 - \lambda^2)x - \frac{\lambda}{2}(1 - \lambda^2) \end{aligned}$$

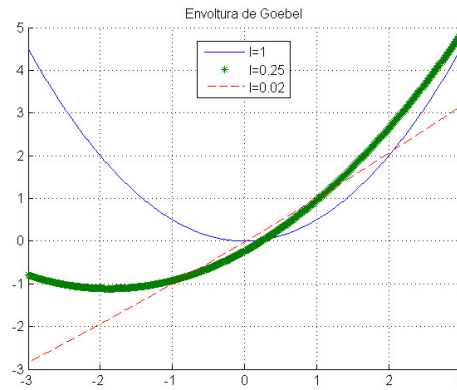


Figura 2.6: $G_{\lambda}f(x) = \frac{\lambda}{2}x^2 + (1 - \lambda^2)x - \frac{\lambda}{2}(1 - \lambda^2)$

En este caso por verificación directa se muestra que *i*) ni *ii*) del Teorema 2.6 se cumplen.

Teorema 2.11.

Dada una función $f \in \Gamma_0(\mathbb{X})$ y $\lambda \in \langle 0, 1 \rangle$. La envoltura de Goebel $G_\lambda f$ es diferenciable con gradiente $\nabla G_\lambda f = \frac{1-\lambda^2}{\lambda}(x - P_\lambda f(x)) + \lambda x$

Demostración.

Vea [5] Lema 2.3 □

Teorema 2.12.

Para $0 < \beta \leq 1$, tenemos que

$$\begin{aligned} G_\beta f &= (1 + \beta)\mathcal{P}(f, 1 - \beta, 0, \beta) + \beta\mathcal{Q} \\ (2.5) \quad G_\beta f(x) &= (1 + \beta)^2 \mathcal{P}\left(f, \frac{1 - \beta}{1 + \beta}, \mathcal{Q}, \frac{2\beta}{1 + \beta}\right)\left(\frac{x}{1 + \beta}\right) \end{aligned}$$

Demostración.

Por la definición de la envoltura de Moreau

$$\begin{aligned} G_\beta f(x) &= (1 - \beta^2)e_\beta f(x) + \beta\mathcal{Q}(x) \\ &= (1 - \beta^2) \left(f \# \frac{1}{\beta} \mathcal{Q} \right) (x) + \beta\mathcal{Q}(x) \\ &= (1 - \beta^2) \inf_w \left\{ f(w) + \frac{1}{\beta} \mathcal{Q}(x - w) \right\} + \beta\mathcal{Q}(x) \\ (2.6) \quad &= (1 - \beta^2) \inf_u \left\{ f\left(\frac{u}{1 - \beta}\right) + \frac{1}{\beta} \mathcal{Q}\left(x - \frac{u}{1 - \beta}\right) + \frac{\beta}{1 - \beta^2} \mathcal{Q}(x) \right\} \end{aligned}$$

usando el hecho que

$$\frac{1}{(1 - \beta)^2} + \frac{1}{\beta(1 - \beta)} = \frac{1}{\beta(1 - \beta)^2} \text{ y } \frac{1}{\beta(1 - \beta)} - \frac{1}{(1 - \beta^2)} = \frac{1}{\beta(1 - \beta)^2}$$

tenemos

$$\begin{aligned} &\frac{1}{\beta} \mathcal{Q}\left(x - \frac{u}{1 - \beta}\right) + \frac{\beta}{1 - \beta^2} \mathcal{Q}(x) = \\ &= \frac{1}{2\beta} \left\| x - \frac{u}{1 - \beta} \right\|^2 + \frac{\beta}{1 - \beta^2} \frac{\|u\|^2}{2} \\ &= \frac{1}{2\beta} \left(\frac{\|u\|^2}{(1 - \beta)^2} - 2 \left\langle x, \frac{u}{1 - \beta} \right\rangle + \|x\|^2 \right) + \left(\frac{1}{\beta(1 - \beta^2)} - \frac{1}{\beta} \right) \frac{\|u\|^2}{2} \\ &= \frac{1}{\beta(1 - \beta)^2} \frac{\|u\|^2}{2} - \frac{1}{\beta(1 - \beta)} \langle x, u \rangle + \frac{1}{\beta(1 - \beta^2)} \frac{\|u\|^2}{2} \\ &= \frac{1}{(1 - \beta)^2} \frac{\|u\|^2}{2} + \frac{1}{\beta(1 - \beta)} \frac{\|x\|^2}{2} - \frac{1}{\beta(1 - \beta)} \langle x, u \rangle + \frac{1}{\beta(1 - \beta)} \frac{\|u\|^2}{2} - \\ &\quad - \frac{1}{1 - \beta^2} \frac{\|u\|^2}{2} \\ &= \mathcal{Q}\left(\frac{u}{1 - \beta}\right) + \frac{1}{\beta(1 - \beta)} \mathcal{Q}(x - u) - \frac{1}{1 - \beta^2} \mathcal{Q}(x) \end{aligned}$$

Reemplazamos esto en la ecuación 2.6 , obtenemos

$$\begin{aligned}
G_\beta f(x) &= (1 - \beta^2) \inf_u \left\{ f\left(\frac{u}{1 - \beta}\right) + \mathcal{Q}\left(\frac{u}{1 - \beta}\right) + \frac{\mathcal{Q}(x - u)}{\beta(1 - \beta)} - \frac{\mathcal{Q}(x)}{1 - \beta^2} \right\} \\
&= (1 + \beta) \inf_u \left\{ (1 - \beta)f\left(\frac{u}{1 - \beta}\right) + 1 - \beta\mathcal{Q}\left(\frac{u}{1 - \beta}\right) + \frac{1}{\beta}\mathcal{Q}(x - u) \right\} - \mathcal{Q}(x) \\
&= (1 + \beta) \inf_u \left\{ (1 - \beta)(f + \mathcal{Q})\left(\frac{u}{1 - \beta}\right) + \frac{1}{\beta}\mathcal{Q}(x - u) \right\} - \mathcal{Q}(x)
\end{aligned}$$

Usando la Definición 1.28

$$\begin{aligned}
G_\beta f(x) &= (1 + \beta) \left[\{(1 + \beta)(f + \mathcal{Q})\left(\frac{\cdot}{1 - \beta}\right)\} \# \frac{1}{\beta} \mathcal{Q} \right] (x) - \mathcal{Q}(x) \\
&= (1 + \beta) [\{(1 + \beta)(f + \mathcal{Q})^*\}^* \# (\beta \mathcal{Q})^*] (x) - \mathcal{Q}(x)
\end{aligned}$$

De donde

$$G_\beta f(x) = (1 + \beta) [(1 + \beta)(f + \mathcal{Q})^* + \beta \mathcal{Q}]^* (x) - \mathcal{Q}(x)$$

A continuación, aplicamos el Teorema 1.23 (1) , el Teorema 1.17 , Ejemplo 1.16 y el Teorema 1.22 (1) obtenemos la primera igualdad en la ecuación 2.5.

$$\begin{aligned}
G_\beta f(x) &= (1 + \beta) [((1 - \beta)(f + g)^* + \beta \mathcal{Q})^* (x) - \mathcal{Q}(x)] + \beta \mathcal{Q}(x) \\
&= (1 + \beta) [(1 - \beta)(f + g)^* + \beta \mathcal{Q}]^* (x) - \mathcal{Q}(x) \\
&= (1 + \beta)^2 \left(\frac{1 - \beta}{1 + \beta} (f + g)^* + \frac{\beta}{1 + \beta} \mathcal{Q} \right)^* \left(\frac{x}{1 + \beta} \right) - \mathcal{Q}(x) \\
&= (1 + \beta)^2 \left[\left(\frac{1 - \beta}{1 + \beta} (f + g)^* + \frac{2\beta}{1 + \beta} (\mathcal{Q} + \mathcal{Q})^* \right)^* \left(\frac{x}{1 + \beta} \right) - \frac{1}{2} \frac{\|x\|^2}{(1 + \beta)^2} \right]
\end{aligned}$$

Para cada x , por la definición de media proximal, por el Teorema 1.22 (1) y el Ejemplo 1.16 se tiene la segunda parte de la igualdad de 2.5.

$$G_\beta f(x) = (1 + \beta)^2 \mathcal{P} \left(f, \frac{1 - \beta}{1 + \beta}, \mathcal{Q}, \frac{2\beta}{1 + \beta} \right) \left(\frac{x}{1 + \beta} \right)$$

□

2.2. Convergencia de la media proximal en presencia de errores de computo

Para los siguientes resultados de este capítulo, vamos a considerar $f \in \Gamma_0(\mathbb{X})$ tal que

$$(2.7) \quad \text{dom}(f) = \{x \in \mathbb{X} / f(x) < \infty\} \neq \emptyset, \quad f(x) \geq -a \text{ para todo } x \in \mathbb{X}, \text{ y}$$

$$(2.8) \quad \lim_{\|x\| \rightarrow \infty} f(x) = \infty$$

por lo tanto

$$(2.9) \quad \arg \min(f) = \{z \in \mathbb{X} / f(z) = \inf(f)\} \neq \emptyset$$

entonces escogemos

$$(2.10) \quad x^* \in \arg \min(f)$$

Tomamos un número entero positivo M tal que

$$(2.11) \quad M > \inf(f) + 4$$

y por 2.8 se tiene $M_0 > 1$ tal que

$$(2.12) \quad f(z) > M + 4, \text{ para todo } z \in \mathbb{X}, \text{ satisfaciendo } \|z\| \geq M_0 - 1$$

con lo anterior se tiene que

$$(2.13) \quad \|x^*\| < M_0 - 1$$

además asumiremos que

$$(2.14) \quad 0 < a_1 < a_2 \leq \frac{M_0^{-2}}{2}$$

Lema 2.13.

Asumiendo que se cumple

$$(2.15) \quad \lambda_k \in [a_1, a_2], k = 0, 1, 2, \dots$$

y que la sucesión $\{x_k\}_{k=0}^\infty$ satisface las condiciones

$$(2.16) \quad f(x_0) \leq M$$

$$(2.17) \quad f(x_{k+1}) + 2^{-1}\lambda_k\|x_{k+1} - x_k\|^2 \leq \inf_{u \in \mathbb{X}} (f(u) + 2^{-1}\lambda_k\|u - x_k\|^2) + 1$$

entonces $\|x_k\| \leq M_0$, para todo $k = 0, 1, \dots$

Demostración. Vamos a realizar la prueba por inducción, de la relación 2.16 se tiene que $f(x_0) \leq M$ luego de la relación 2.12 se tiene que

$$(2.18) \quad \|x_0\| \leq M_0$$

Ahora supongamos que para $k \in \mathbb{Z}_0^+$ se tiene que

$$(2.19) \quad \|x_k\| \leq M_0$$

recordemos que

$$\|x^*\| < M_0 - 1, \quad 0 < a_1 < a_2 \leq \frac{1}{2M_0^2}, \quad x^* \in \arg \min(f) \text{ y } M > \inf(f) + 4$$

entonces se tiene que

$$\begin{aligned} \|x^* - x_k\| &\leq \|x^*\| + \|x_k\| \leq M_0 + M_0 = 2M_0, \quad \lambda_k \in [a_1, a_2] \\ 2^{-1}\lambda_k\|x^* - x_k\|^2 &\leq 2^{-1}\lambda_k(2M_0)^2 \leq 1, \quad \text{pues } \lambda_k \leq \frac{1}{2M_0^2} \\ 2^{-1}\lambda_k\|x^* - x_k\|^2 &\leq 1 \end{aligned}$$

Así que de la relación 2.17 se tiene que

$$\begin{aligned}
 f(x_{k+1}) + 2^{-1}\lambda_k \|x_{k+1} - x_k\|^2 &\leq \inf_{u \in \mathbb{X}} (f(u) + 2^{-1}\lambda_k \|u - x_k\|^2) + 1 \\
 &\leq f(x^*) + 2^{-1}\lambda_k \|x^* - x_k\|^2 + 1 \\
 &\leq f(x^*) + 2 \\
 f(x_{k+1}) &\leq f(x_{k+1}) + 2^{-1}\lambda_k \|x_{k+1} - x_k\|^2 \leq f(x^*) + 2 \\
 f(x_{k+1}) &\leq \inf(f) + 2 < M, \text{ de 2.11} \\
 f(x_{k+1}) &\leq M \\
 (2.20) \quad \|x_{k+1}\| &\leq M_0, \text{ de 2.12}
 \end{aligned}$$

Por lo tanto, de las relaciones 2.18 , 2.19 y 2.20 se ha probado que $\|x_{k+1}\| \leq M_0$, para todo $k \in \mathbb{Z}_0^+$. □

Lema 2.14.

$$(2.21) \quad \text{Sean } \lambda_k \in [a_1, a_2], \quad \varepsilon_k \in \langle 0, 1 \rangle, \quad k = 0, 1, 2, \dots$$

y la sucesión $\{x_k\}_{k=0} \subseteq \mathbb{X}$ satisfaciendo

$$(2.22) \quad f(x_0) \leq M$$

y para todo $k = 0, 1, \dots$ se cumple que

$$(2.23) \quad f(x_{k+1}) + 2^{-1}\lambda_k \|x_{k+1} - x_k\|^2 \leq \inf_{u \in \mathbb{X}} (f(u) + 2^{-1}\lambda_k \|u - x_k\|^2) + \varepsilon_k$$

Entonces se cumple que

$$(2.24) \quad (2/\lambda_k)[f(x_{k+1}) - f(x^*)] + \|x_{k+1} - x_k\|^2 \leq 2\varepsilon_k a_1^{-1} + \|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 + 8M_0(\varepsilon_k a_1^{-1})^{1/2}$$

Además para todo par de naturales tales que $m > n$ se tiene que

$$(2.25) \quad \sum_{i=n}^m 2a_1^{-1}[f(x_i) - f(x^*)] + \sum_{i=n}^m \|x_{i-1} - x_i\|^2 \leq 4M_0^2 + \sum_{i=n-1}^{m-1} [2a_1^{-1}\varepsilon_i + 8M_0(\varepsilon_i a_1^{-1})^{1/2}]$$

Demostración. Para realizar esta demostración vamos a tener en cuenta las relaciones 2.22, 2.23 y el Lema 2.13 de aquí se tiene

$$(2.26) \quad \|x_k\| \leq M_0 \text{ para todo } k \in \mathbb{Z}_0^+$$

De la relación 2.23 para todo $k \in \mathbb{Z}_0^+$ se tiene

$$\begin{aligned}
 f(x_{k+1}) &\leq f(x_{k+1}) + 2^{-1}\lambda_k \|x_{k+1} - x_k\|^2 \\
 &\leq \inf_{u \in \mathbb{X}} \{f(u) + 2^{-1}\lambda_k \|u - x_k\|^2\} + \varepsilon_k \\
 &\quad \text{tome } u = x_k \in \mathbb{X} \\
 &\leq f(x_k) + 2^{-1}\lambda_k \|x_k - x_k\|^2 + \varepsilon_k \\
 (2.27) \quad f(x_{k+1}) &\leq f(x_k) + \varepsilon_k
 \end{aligned}$$

Probaremos la relación 2.24; para ello sea $k \in \mathbb{Z}_0^+$ existe $y_{k+1} \in \mathbb{X}$ y de la relación 2.23 se tiene que

$$(2.28) \quad f(y_{k+1}) + 2^{-1}\lambda_k\|y_{k+1} - x_k\|^2 \leq f(x) + 2^{-1}\lambda_k\|x - x_k\|^2, \quad \forall x \in \mathbb{X}$$

Estimamos $\|x_{k+1} - y_{k+1}\|$, tomamos

$$(2.29) \quad z = 2^{-1}(x_{k+1} + y_{k+1})$$

tenemos:

$$\begin{aligned} & 2^{-1}\|y_{k+1} - x_k\|^2 + 2^{-1}\|x_{k+1} - x_k\|^2 - \|2^{-1}(x_{k+1} + y_{k+1}) - x_k\|^2 = 2^{-1}\|y_{k+1}\|^2 + \\ & \quad + 2^{-1}\|x_k\|^2 - \langle y_{k+1}, x_k \rangle + 2^{-1}\|x_{k+1}\|^2 + 2^{-1}\|x_k\|^2 - \langle x_{k+1}, x_k \rangle \\ & \quad - \|2^{-1}(y_{k+1} + x_{k+1})\|^2 - \|x_k\|^2 + \langle x_k, x_{k+1} + y_{k+1} \rangle \\ & = 2^{-1}\|y_{k+1}\|^2 + 2^{-1}\|x_{k+1}\|^2 - \|2^{-1}(y_{k+1} + x_{k+1})\|^2 \\ & = 2^{-1}\|y_{k+1}\|^2 + 2^{-1}\|x_{k+1}\|^2 - \left\{ \frac{\|y_{k+1}\|^2}{4} + \frac{\|x_{k+1}\|^2}{4} + \langle y_{k+1}, x_{k+1} \rangle \right\} \\ & = \frac{\|y_{k+1}\|^2}{4} + \frac{\|x_{k+1}\|^2}{4} + \langle y_{k+1}, x_{k+1} \rangle \end{aligned}$$

$$2^{-1}\|y_{k+1} - x_k\|^2 + 2^{-1}\|x_{k+1} - x_k\|^2 - \|2^{-1}(x_{k+1} + y_{k+1}) - x_k\|^2 = \|2^{-1}(y_{k+1} - x_{k+1})\|^2$$

(2.30)

$$2^{-1}\|y_{k+1} - x_k\|^2 + 2^{-1}\|x_{k+1} - x_k\|^2 - \|2^{-1}(y_{k+1} - x_{k+1})\|^2 = \|2^{-1}(x_{k+1} + y_{k+1}) - x_k\|^2$$

Ahora de las relaciones 2.29, 2.30 y la convexidad de f y se tiene

$$\begin{aligned} f(z) + 2^{-1}\lambda_k\|z - x_k\|^2 &= f\left(\frac{x_{k+1}}{2} + \frac{y_{k+1}}{2}\right) + 2^{-1}\lambda_k\|z - x_k\|^2 \\ &\leq \frac{1}{2}f(x_{k+1}) + \frac{1}{2}f(y_{k+1}) + 2^{-1}\lambda_k\|2^{-1}(x_{k+1} + y_{k+1}) - x_k\|^2 \\ &\leq \frac{1}{2}\{f(x_{k+1}) + f(y_{k+1})\} + 2^{-1}\lambda_k\{2^{-1}\|y_{k+1} - x_k\|^2 + \\ &\quad 2^{-1}\|x_{k+1} - x_k\|^2 - \|2^{-1}(y_{k+1} - x_{k+1})\|^2\} \\ &\leq \frac{1}{2}\{f(x_{k+1}) + 2^{-1}\lambda_k\|x_{k+1} - x_k\|^2\} + \\ &\quad + \frac{1}{2}\{f(y_{k+1}) + 2^{-1}\lambda_k\|y_{k+1} - x_k\|^2\} + \\ &\quad + \frac{1}{2}\lambda_k\{-\|2^{-1}(y_{k+1} - x_{k+1})\|^2\} \text{ de la relación 2.23} \\ &\leq \inf_{x \in \mathbb{X}} \{f(x) + 2^{-1}\lambda_k\|x - x_k\|^2\} + \frac{\varepsilon_k}{2} - \\ &\quad - 2^{-1}\lambda_k\{\|2^{-1}(y_{k+1} - x_{k+1})\|^2\} \\ &\leq \inf_{x \in \mathbb{X}} \{f(x) + 2^{-1}\lambda_k\|x - x_k\|^2\} + \frac{\varepsilon_k}{2} + \\ &\quad + 2^{-1}\lambda_k\{\|2^{-1}(y_{k+1} - x_{k+1})\|^2\} \end{aligned}$$

Además de 2.15 tenemos $\lambda_k \in [a_1, a_2]$; $0 < a_1 < \lambda_k < a_2 \leq \frac{M_0^{-1}}{2}$

$$\begin{aligned}
 2^{-1}\lambda_k \|2^{-1}(y_{k+1} - x_{k+1})\|^2 &\leq \frac{\varepsilon_k}{2} \\
 \lambda_k \|2^{-1}(y_{k+1} - x_{k+1})\|^2 &\leq \varepsilon_k \\
 \|2^{-1}(y_{k+1} - x_{k+1})\|^2 &\leq \frac{\varepsilon_k}{\lambda_k} < \frac{\varepsilon_k}{a_1} \\
 \|2^{-1}(y_{k+1} - x_{k+1})\| &\leq (\varepsilon_k a_1^{-1})^{\frac{1}{2}} \\
 (2.31) \quad \|y_{k+1} - x_{k+1}\| &\leq 2(\varepsilon_k a_1^{-1})^{\frac{1}{2}}
 \end{aligned}$$

Ahora estimamos $f(x^*) - f(x_{k+1})$

De 2.28 se tiene que

$$f(y_{k+1}) + 2^{-1}\lambda_k \|y_k - x_k\|^2 \leq f(x) + 2^{-1}\lambda_k \|x - x_k\|^2$$

$0 \in \partial f(y_{k+1}) + \lambda_k(y_{k+1} - x_k)$, y para todo $u \in \mathbb{X}$

$$(2.32) \quad f(u) - f(y_{k+1}) \geq \lambda_k \langle x_k - y_{k+1}, u - y_{k+1} \rangle$$

Ahora tomando $u = x^*$, se tiene

$$(2.33) \quad f(x^*) - f(y_{k+1}) \geq \lambda_k \langle x_k - y_{k+1}, x^* - y_{k+1} \rangle$$

de 2.23 se tiene que

$$f(x_{k+1}) + 2^{-1}\lambda_k \|x_{k+1} - x_k\|^2 \leq f(y_{k+1}) + 2^{-1}\lambda_k \|y_{k+1} - x_k\|^2 + \varepsilon_k$$

$$(2.34) \quad f(y_{k+1}) - f(x_{k+1}) \geq 2^{-1}\lambda_k (\|x_{k+1} - x_k\|^2 - \|y_{k+1} - x_k\|^2) - \varepsilon_k$$

Ahora usando 2.33 se tiene

$$\begin{aligned}
 f(x^*) - f(x_{k+1}) &= f(x^*) - f(y_{k+1}) + f(y_{k+1}) - f(x_{k+1}) \\
 &\geq \lambda_k \langle x_k - y_{k+1}; x^* - y_{k+1} \rangle + \\
 &\quad + 2^{-1}\lambda_k (\|x_{k+1} - x_k\|^2 - \|y_{k+1} - x_k\|^2) - \varepsilon_k
 \end{aligned}$$

Observación 2.3.

$$\begin{aligned}
 \langle x_k - y_{k+1}; x^* - y_{k+1} \rangle &= \langle x_k - x^* + x^* - y_{k+1}; x^* - y_{k+1} \rangle \\
 &= \langle x_k - x^*; x^* - y_{k+1} \rangle + \|x^* - y_{k+1}\|^2 \\
 &= \langle x_k - x^*; \underline{x^* - x_k} + \underline{x_k - y_{k+1}} \rangle + \|x^* - y_{k+1}\|^2 \\
 &= -\|x_k - x^*\|^2 + \langle x_k - x^*; x_k - y_{k+1} \rangle + \|x^* - y_{k+1}\|^2 \\
 &= -\|x_k - x^*\|^2 + \langle x_k - y_{k+1} + y_{k+1} - x^*; x_k - y_{k+1} \rangle + \\
 &\quad + \|x^* - y_{k+1}\|^2 \\
 &= -\|x_k - x^*\|^2 + \|x_k - y_{k+1}\|^2 + \langle y_{k+1} - x^*, x_k - y_{k+1} \rangle + \\
 &\quad + \|x^* - y_{k+1}\|^2 \\
 &= -\|x_k - x^*\|^2 + \|x_k - y_{k+1}\|^2 - \langle x^* - y_{k+1}, x_k - y_{k+1} \rangle + \\
 &\quad + \|x^* - y_{k+1}\|^2 \\
 \langle x_k - y_{k+1}; x^* - y_{k+1} \rangle &= -\|x_k - x^*\|^2 + \|x_k - y_{k+1}\|^2 + \|x^* - y_{k+1}\|^2
 \end{aligned}$$

Por lo tanto de la observación se tiene

$$\begin{aligned}
 f(x^*) - f(x_{k+1}) &\geq 2^{-1}\lambda_k(\|x_k - y_{k+1}\|^2 + \|x^* - y_{k+1}\|^2 - \|x_k - x^*\|^2) + \\
 &\quad + 2^{-1}\lambda_k(\|x_{k+1} - x_k\|^2 - \|y_{k+1} - x_k\|^2) - \varepsilon_k \\
 (2.35) \quad &\geq 2^{-1}\lambda_k(\|y_{k+1} - x^*\|^2 - \|x_k - x^*\|^2 + \|x_k - x_{k+1}\|^2) - \varepsilon_k
 \end{aligned}$$

Por lo tanto de 2.28, 2.21, 2.26, 2.13, 2.14, 2.11 y para un entero $q \geq 1$

$$\begin{aligned}
 f(y_{k+1}) + 2^{-1}\lambda_k\|y_{k+1} - x_k\|^2 &\leq f(x) + 2^{-1}\lambda_k\|x - x_k\|^2 \\
 &\quad \text{para todo } x \in \mathbb{X}, \text{ tomo } x = x^*, \lambda_k \leq a_2 \\
 f(y_{k+1}) &\leq f(x^*) + 2^{-1}a_2\|x^* - x_k\|^2 - \|y_{k+1} - x_k\|^2 \\
 f(y_{k+1}) &\leq f(x^*) + 2^{-1}a_2\|x^* - y_k\|^2 \\
 f(y_q) &\leq f(x^*) + 2^{-1}a_2\|x^* - y_q\|^2 \leq f(x^*) + 1 \\
 f(y_q) &\leq M
 \end{aligned}$$

La desigualdad 2.12 implica que $\|y_q\| \leq M_0$, $q = 1, 2, 3, \dots$

$$(2.36) \quad \|y_q\| \leq M_0, \quad q = 1, 2, 3, \dots$$

Ahora usamos 2.35 y 2.36 y obtenemos la estimación de $f(x^*) - f(x_{k+1})$ sin términos que contengan y_{k+1} , por 2.26 y 2.31

$$\begin{aligned}
 \|x_k - y_{k+1}\|^2 &= \|(x - k - x_{k+1}) - (y_{k+1} - x_{k+1})\|^2 \\
 &= \|(x - k - x_{k+1})\|^2 + \|y_{k+1} - x_{k+1}\|^2 - \langle x_k - x_{k+1}, y_{k+1} - x_{k+1} \rangle \\
 &\geq \|(x - k - x_{k+1})\|^2 + \|y_{k+1} - x_{k+1}\|^2 - 2\|x_k - x_{k+1}\|\|y_{k+1} - x_{k+1}\| \\
 &\geq \|(x - k - x_{k+1})\|^2 - 2\|x_k - x_{k+1}\|\|y_{k+1} - x_{k+1}\|
 \end{aligned}$$

Observación 2.4.

De 2.26

$$\begin{aligned}
 \|x_k\| &\leq M_0, \quad \text{para todo } k \in \mathbb{Z}^+ \\
 \|x_k - x_{k+1}\| &\leq \|x_k\| + \|x_{k+1}\| \leq 2M_0 \\
 2\|x_k - x_{k+1}\| &\leq \|x_k\| + \|x_{k+1}\| \leq 4M_0 \\
 \|y_{k+1} - x_{k+1}\| &\leq 2(\varepsilon_k a_1^{-1})^{\frac{1}{2}}
 \end{aligned}$$

de estas dos ultimas relaciones se tiene que

$$2\|x_k - x_{k+1}\|\|y_{k+1} - x_{k+1}\| \leq 8M_0(\varepsilon_k a_1^{-1})^{\frac{1}{2}}$$

Por lo tanto de la observación, se tiene

$$(2.37) \quad \|x_k - y_{k+1}\|^2 \geq \|x_k - x_{k+1}\|^2 - 8M_0(\varepsilon_k a_1^{-1})^{\frac{1}{2}}$$

y de 2.26, 2.13 y 2.31 tenemos

$$\begin{aligned}
 \|y_{k+1} - x^*\|^2 &= \|(x_{k+1} - x^*) + (y_{k+1} - x_{k+1})\|^2 \\
 &= \|x_{k+1} - x^*\|^2 + \|y_{k+1} - x_{k+1}\|^2 + 2\langle x_{k+1} - x^*, y_{k+1} - x_{k+1} \rangle \\
 &\geq \|x_{k+1} - x^*\|^2 + \|y_{k+1} - x_{k+1}\|^2 + 2\|x_{k+1} - x^*\|\|y_{k+1} - x_{k+1}\| \\
 (2.38) \quad &\geq \|x_{k+1} - x^*\|^2 - 8M_0(\varepsilon_k a_1^{-1})^{\frac{1}{2}}
 \end{aligned}$$

De 2.35 y 2.38 tenemos

$$\begin{aligned}
f(x^*) - f(x_{k+1}) &\geq 2^{-1}\lambda_k\{\|y_{k+1} - x^*\|^2 - \|x_k - x^*\|^2 + \|x_k - x_{k+1}\|^2\} - \varepsilon_k \\
&\geq 2^{-1}\lambda_k\{\|x_{k+1} - x^*\|^2 - 8M_0(\varepsilon_k a_1^{-1})^{\frac{1}{2}} - \|x^* - x_k\|^2 + \\
&\quad + \|x_k - x_{k+1}\|^2\} - \varepsilon_k \\
f(x_{k+1}) - f(x^*) &\leq 2^{-1}\lambda_k\{-\|x_{k+1} - x^*\|^2 + 8M_0(\varepsilon_k a_1^{-1})^{\frac{1}{2}} + \|x^* - x_k\|^2 - \\
&\quad - \|x_k - x_{k+1}\|^2\} + \varepsilon_k
\end{aligned}$$

$$\begin{aligned}
f(x_{k+1}) - f(x^*) + 2^{-1}\lambda_k\|x_k - x_{k+1}\|^2 &\leq \varepsilon_k + 2^{-1}\lambda_k\{\|x^* - x_k\|^2 - \|x_{k+1} - x^*\|^2\} + \\
&\quad + 2^{-1}\lambda_k 8M_0(\varepsilon_k a_1^{-1})^{\frac{1}{2}}
\end{aligned}$$

De 2.21 se tiene $a_1^{-1} \geq \lambda_k^{-1} \geq a_2^{-1} \geq 2M_0^2$. Por lo tanto

$$\frac{2}{\lambda_k}\{f(x_{k+1}) - f(x^*)\} + \|x_{k+1} - x_k\|^2 \leq 2\varepsilon_k a_1^{-1} + \|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 + 8M_0(\varepsilon_k a_1^{-1})^{1/2}$$

Así queda probada la afirmación 2.24. Ahora probaremos la afirmación 2.25, Por 2.24, $2\lambda_k^{-1} \geq \frac{2}{a_2}$, 2.21, 2.13, 2.26, y considerando $m > n$ un par de números naturales

$$\frac{2}{a_2}\{f(x_{k+1}) - f(x^*)\} + \|x_{k+1} - x_k\|^2 \leq 2\varepsilon_k a_1^{-1} + \|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 + 8M_0(\varepsilon_k a_1^{-1})^{1/2}$$

tomamos sumatoria y aplicamos la suma telescópica tenemos:

$$\begin{aligned}
\sum_{i=n}^m \frac{2}{a_2}\{f(x_i) - f(x^*)\} + \sum_{i=n}^m \|x_{i+1} - x_i\|^2 &\leq \|x_{n-1} - x^*\|^2 + \\
&\quad + \sum_{i=n-1}^{m-1} \{2\varepsilon_i a_1^{-1} + 8M_0(\varepsilon_i a_1^{-1})^{1/2}\} \\
&\leq 4M_0^2 + \sum_{i=n-1}^{m-1} \{2\varepsilon_i a_1^{-1} + 8M_0(\varepsilon_i a_1^{-1})^{1/2}\}
\end{aligned}$$

Así se completa la prueba. □

Teorema 2.15.

Sean

$$(2.39) \quad \lambda_k \in [a_1, a_2], \quad k = 0, 1, 2, \dots$$

$$(2.40) \quad \Delta \in \langle 0, 1], \quad L \in \mathbb{N} \text{ tales que } L > 2(4M_0^2 + 1)a_2\Delta^{-1}$$

$$(2.41) \quad \varepsilon^{1/2}(L+1)(2a_1^{-1} + 8M_0a_1^{-1/2}) \leq 1 \text{ y } \varepsilon(L+1) \leq \frac{\Delta}{4}$$

Asumimos que la sucesión $\{x_k\}_k \subseteq \mathbb{X}$ satisface

$$(2.42) \quad f(x_0) \leq M$$

y también para todo $k \in \mathbb{Z}_0^+$

$$(2.43) \quad f(x_{k+1}) + 2^{-1}\lambda_k\|x_{k+1} - x_k\|^2 \leq \inf(f + 2^{-1}\lambda_k\|\cdot - x_k\|^2) + \varepsilon,$$

entonces para todo $k > L$ se tiene que $f(x_k) \leq \inf(f) + \Delta$

Demostración. Recordemos las hipótesis 2.39, 2.40, 2.41, 2.42, 2.43 y el Lema 2.14, aplicados a $n \in \mathbb{N}$ y $m = n + l$

$$\begin{aligned} \sum_{i=n}^m \frac{2}{a_2} \{f(x_i) - f(x^*)\} + \sum_{i=n}^m \|x_{i+1} - x_i\|^2 &\leq 4M_0^2 + \\ &+ \sum_{i=n-1}^{m-1} \{2\varepsilon_i a_1^{-1} + 8M_0(\varepsilon_i a_1^{-1})^{1/2}\} \end{aligned}$$

tomamos, $\varepsilon = \max\{\varepsilon_i\}$

$$\begin{aligned} \sum_{i=n}^{n+l} \frac{2}{a_2} \{f(x_i) - f(x^*)\} + \sum_{i=n}^{n+l} \|x_{i+1} - x_i\|^2 &\leq 4M_0^2 + \\ &+ \sum_{i=n-1}^{n+l-1} \{2\varepsilon^{1/2} a_1^{-1} + 8M_0(\varepsilon a_1^{-1})^{1/2}\} \\ &\leq 4M_0^2 + \\ &+ (l+1)\varepsilon^{1/2} \sum_{i=n-1}^{n+l-1} \{2a_1^{-1} + 8M_0(a_1^{-1})^{1/2}\} \\ \sum_{i=n}^m \frac{2}{a_2} \{f(x_i) - f(x^*)\} &\leq 4M_0^2 + \\ &+ (l+1)\varepsilon^{1/2} \sum_{i=n-1}^{n+l-1} \{2a_1^{-1} + 8M_0(a_1^{-1})^{1/2}\} \\ &\text{de 2.41, se tiene} \end{aligned}$$

$$(2.44) \quad \sum_{i=n}^m \frac{2}{a_2} \{f(x_i) - f(x^*)\} \leq 4M_0^2 + 1$$

tomamos el $\min\{f(x_i) - f(x^*) \mid i = n, n+1, \dots, n+l\}$ y en 2.44 tenemos

$$\begin{aligned} (l+1)2a_2^{-1} \min\{f(x_i) - f(x^*) \mid i = n, n+1, \dots, n+l\} &\leq \sum_{i=n}^{n+l} \frac{2}{a_2} \{f(x_i) - f(x^*)\} \\ &\leq 4M_0^2 + 1 \end{aligned}$$

$$\begin{aligned} \min_i \{f(x_i) - f(x^*) \mid i = n, \dots, n+l\} &\leq (4M_0^2 + 1)(l+1)^{-1} 2^{-1} a_2 \\ (2.45) \quad &< \frac{\Delta}{4} \end{aligned}$$

lo cual se cumple pues de 2.40 tenemos $2(4m_0^2 + 1)a_2\Delta^{-1} < L$ entonces $2^{-1}(4m_0^2 + 1)a_2 < \frac{L\Delta}{4}$ de aquí $2^{-1}(4m_0^2 + 1)a_2(l+1)^{-1} < \frac{L\Delta}{4(l+1)} < \frac{\Delta}{4}$.

ahora como 2.45 es valido para todo $n \in \mathbb{N}$, existe una sucesión estrictamente creciente de números naturales $\{s_i\}_{i=1}^\infty$ tal que

$$(2.46) \quad s_1 \in \{1, \dots, 1+L\}, \quad s_{i+1} - s_i \in [1, 1+L], \quad i = 1, \dots$$

$$(2.47) \quad f(x_{s_1}) - f(x^*) \leq \frac{\Delta}{4}, \quad i = 1, \dots$$

Sea un entero $j \geq L+1$ por (2.46) existe un número natural i tal que $s_i \leq j \leq s_{i+1}$ de aquí $0 \leq j - s_i \leq s_{i+1} - s_i \in [1, L+1]$ entonces

$$(2.48) \quad 0 \leq j - s_i \leq L+1$$

en 2.9 para todo $k \in \mathbb{Z}_0^+$,

$$f(x_{k+1}) + 2^{-1}\lambda k \|x_{k+1} - x_k\|^2 \leq \inf_{u \in \mathbb{X}} (f(u) + 2^{-1}\lambda k \|u - x_k\|^2) + \epsilon$$

de aquí se tiene

$$\begin{aligned} f(x_{k+1}) &\leq \inf_{u \in \mathbb{X}} (f(u) + 2^{-1}\lambda k \|u - x_k\|^2) + \epsilon \\ &\leq f(u) + 2^{-1}\lambda k \|u - x_k\|^2 + \epsilon, \quad \text{tomemos } u = x_k \\ &\leq f(x_k) + 2^{-1}\lambda k \|x_k - x_k\|^2 + \epsilon \\ (2.49) \quad f(x_{k+1}) &\leq f(x_k) + \epsilon \end{aligned}$$

Usando 2.48, 2.47 y 2.7 $f(x_{s_i}) \leq f(x^*) + \frac{\Delta}{4}$, $\varepsilon(t+1) \leq \frac{\Delta}{4}$ se tiene que

$$\begin{aligned} f(x_j) &\leq f(x_{s_i}) + \epsilon(L+1) \\ &\leq (f(x^*) + \frac{\Delta}{4}) + \frac{\Delta}{4} \\ &\leq f(x^*) + \frac{\Delta}{2} \\ f(x_j) &\leq f(x^*) + \Delta \end{aligned}$$

Así se tiene que $f(x_k) \leq \inf(f) + \Delta$ y el teorema está probado. □

Teorema 2.16.

Sean

$$(2.50) \quad \lambda_k \in [a_1, a_2], \quad k = 0, 1, 2, \dots$$

$L \in \mathbb{N}$ tal que

$$(2.51) \quad L > (4M_0^{-1} + 1)2a_2$$

y $\bar{\varepsilon} \in \mathbb{R}^+$ satisfaciendo

$$(2.52) \quad \bar{\varepsilon}^{1/2}(L+1)(2a_1^{-1} + 8M_0a_1^{-1/2}) \leq 1 \quad y \quad \bar{\varepsilon}(L+1) \leq \frac{1}{4}$$

asumimos que

$$(2.53) \quad \{\varepsilon_i\}_{i=0}^\infty \subseteq \langle 0, \bar{\varepsilon} \rangle, \quad \lim_{i \rightarrow 0} \varepsilon_i = 0$$

y $\gamma > 0$, entonces existe $T_0 \in \mathbb{N}$ tal que para cada sucesión $\{x_k\}_{k=0}^\infty \subseteq \mathbb{X}$ satisfaciendo

$$(2.54) \quad f(x_0) \leq M \quad y$$

$$(2.55) \quad f(x_{k+1}) + 2^{-1}\lambda_k\|x_{k+1} - x_k\|^2 \leq \inf(f + 2^{-1}\lambda_k\|\cdot - x_k\|^2) + \varepsilon_k$$

Para todo $k \in \mathbb{Z}_0^+$, se cumple la desigualdad

$$f(x_k) \leq \inf(f) + \gamma, \quad k > T_0$$

Demostración. Para realizar esta demostración, recordemos el Teorema 2.15 se tiene

(P1) Sea la sucesión $\{x_k\}_{k=0}^\infty \subseteq \mathbb{X}$ satisfaciendo $f(x_0) \leq M$

$$f(x_{k+1}) + 2^{-1}\lambda_k\|x_{k+1} - x_k\|^2 \leq \inf_{u \in \mathbb{X}}(f(u) + 2^{-1}\lambda_k\|u - x_k\|^2) + \bar{\varepsilon}, \quad k = 0, 1, \dots$$

entonces $f(x_k) \leq \inf(f) + 1$, para todo $k \in \mathbb{Z}_0^+$, $k > L$

por hipótesis tenemos $\bar{\varepsilon} > 0$, tomamos $\delta \in \langle 0, \bar{\varepsilon} \rangle$ y un $L_0 \in \mathbb{N}$ tales que cumplan las hipótesis del Teorema 2.15 se tiene

(P2) Para cada sucesión $\{y_i\}_{i=0}^\infty \subseteq \mathbb{X}$ que satisface $f(y_0) \leq \inf(f) + 1$ y

$$f(y_{k+1}) + 2^{-1}\lambda_k\|y_{k+1} - y_k\|^2 \leq \inf_{u \in \mathbb{X}}(f(u) + 2^{-1}\lambda_k\|u - y_k\|^2) + \delta, \quad k = 0, 1, \dots$$

tenemos

$$(2.56) \quad f(y_k) \leq \inf(f) + \gamma \quad \text{para todo } k \geq L_0, k = 0, 1, 2, \dots$$

donde γ es dado por el Teorema 2.15.

Por 2.53, para algún $L_1 \in \mathbb{N}$ se tiene que

$$(2.57) \quad \varepsilon_k < \delta, \quad \text{para todo } k \in \mathbb{N}, k \geq L_1$$

escogemos un número natural T_0 tal que

$$(2.58) \quad T_0 > L_0 + L_1 + L$$

asumimos que la sucesión $\{x_i\}_{i=0}^{\infty} \subset \mathbb{X}$ satisface 2.54 y 2.55 , de (P1) , 2.54, 2.55 y 2.53 se tiene que

$$(2.59) \quad f(x_k) \leq \inf(f) + 1 \text{ para cada } k \in \mathbb{Z}, K > L$$

$$(2.60) \quad y_k = x_k + L + L_1$$

por 2.59 y 2.60

$$(2.61) \quad f(y_0) \leq \inf(f) + 1$$

seguimos de 2.55 , 2.60 y 2.57 para todo $k \in \mathbb{Z}_0^+$

$$(2.62) \quad f(y_{k+1}) + 2^{-1}\lambda_k \|y_{k+1} - y_k\|^2 \leq \inf_{u \in \mathbb{X}} (f(u) + 2^{-1}\lambda_k \|u - y_k\|^2) + \delta, \quad k = 0, 1, \dots$$

Por 2.61, 2.62 y (P2) se tiene

$$(2.63) \quad f(y_k) \leq \inf(f) + \gamma, \text{ para todo } k \in \mathbb{Z}_0^+, k \geq L_0$$

junto con 2.60 y 2.58 esta desigualdad implica que

$$f(x_k) \leq \inf(f) + \gamma, \text{ para todo } k \in \mathbb{Z}_0^+, k > T_0$$

Así el teorema está probado. □

Corolario 2.17.

Asumamos todas las hipótesis del Teorema 2.16 y la sucesión $\{x_k\}_{k=0}^{\infty} \subseteq \mathbb{X}$ satisfaciendo las desigualdades 2.54 y 2.55 , para todo $k \in \mathbb{Z}_0^+$, entonces $\lim_{k \rightarrow \infty} f(x_k) = \inf(f)$ y la sucesión $\{x_k\}_{k=0}^{\infty}$ es acotada , además posee una subsucesión débilmente convergente y el limite de cualquier subsucesión débilmente convergente de $\{x_k\}_{k=0}^{\infty}$ es un minimizador de f .

Corolario 2.18.

Sea $f \in \Gamma_0(\mathbb{X})$ se cumple

$$(\lambda_1 f_1^* + \dots + \lambda_n f_n^*)^* \leq \mathcal{P}_\mu(f, \lambda) \leq \lambda_1 f_1 + \dots + \lambda_n f_n$$

Demostración. Para realizar esta demostración recordemos el Teorema 2.1 ii) , tenemos que

$$\mathcal{P}_\mu(f, \lambda)(x) = \frac{1}{\mu} \left(\inf_{\sum \lambda_i y_i = x} \left\{ \sum_i \lambda_i \mu f_i(y_i) + \sum_i \lambda_i q(y_i) \right\} - q(x) \right)$$

Aquí tomamos $y_i = x$ se obtiene que

$$\mathcal{P}_\mu(f, \lambda) \leq \lambda_1 f_1 + \dots + \lambda_n f_n$$

luego, consideramos la igualdad anterior para μ^{-1} y f^* obtenemos que

$$\mathcal{P}_{\mu^{-1}}(f^*, \lambda) \leq \lambda_1 f_1^* + \cdots + \lambda_n f_n^*$$

tomando la conjugada de Fenchel se obtiene

$$\mathcal{P}_\mu(f, \lambda) \geq (\lambda_1 f_1^* + \cdots + \lambda_n f_n^*)^*$$

es decir, obtenemos que

$$(\lambda_1 f_1^* + \cdots + \lambda_n f_n^*)^* \leq \mathcal{P}_\mu(f, \lambda) \leq \lambda_1 f_1 + \cdots + \lambda_n f_n$$

□

Corolario 2.19.

Se cumple

$$\lambda_1 \inf f_1 + \cdots + \lambda_n \inf f_n \leq \inf \mathcal{P}_\mu(f, \lambda) \leq \inf(\lambda_1 f_1 + \cdots + \lambda_n f_n)$$

Demostración. Del Corolario anterior se obtiene que

$$\begin{aligned} (\lambda_1 f_1^* + \cdots + \lambda_n f_n^*)^* &\leq \mathcal{P}_\mu(f, \lambda) \leq \lambda_1 f_1 + \cdots + \lambda_n f_n \\ \lambda_1 f_1 \# \cdots \# \lambda_n f_n &\leq \mathcal{P}_\mu(f, \lambda) \leq \lambda_1 f_1 + \cdots + \lambda_n f_n \\ \inf(\lambda_1 f_1 + \cdots + \lambda_n f_n) &\leq \inf \mathcal{P}_\mu(f, \lambda) \leq \inf(\lambda_1 f_1 + \cdots + \lambda_n f_n) \\ \lambda_1 \inf f_1 + \cdots + \lambda_n \inf f_n &\leq \inf \mathcal{P}_\mu(f, \lambda) \leq \inf(\lambda_1 f_1 + \cdots + \lambda_n f_n) \end{aligned}$$

□

Teorema 2.20.

Si $f \in \Gamma_0(\mathbb{X})$ y $\bigcap_{i; \lambda_i > 0} \arg \min(f_i) \neq \emptyset$ entonces se cumple que

$$\min \mathcal{P}_\mu(f, \lambda) = \sum_{i; \lambda_i > 0} \lambda_i \min f_i \quad y \quad \arg \min \mathcal{P}_\mu(f, \lambda) = \bigcap_{i > 0} P_\mu f_i$$

en consecuencia $\mathcal{P}_\mu(f, \lambda)$ es convergente.

Demostración. Para realizar esta demostración aplicamos el corolario 2.19 y además la hipótesis del Teorema $\bigcap_{i; \lambda_i > 0} \arg \min(f_i) \neq \emptyset$, entonces se tiene

que $\min \mathcal{P}_\mu(f, \lambda) = \sum_{i; \lambda_i > 0} \lambda_i \min f_i$ y por el Teorema 2.6(iii) se cumple que

$$\arg \min \mathcal{P}_\mu(f, \lambda) = \bigcap_{i; \lambda_i > 0} \arg \min(f_i) = \bigcap_{i; \lambda_i > 0} \arg \min e_\mu f_i = \bigcap_{i; \lambda_i > 0} P_\mu f_i \quad \square$$

Capítulo 3

Aplicación y extensión

En éste capítulo presentamos ejemplos de aplicación, así como también una extensión de la media proximal, primero vamos a dar una aplicación donde la media proximal alivia el problema de demostrar que la envoltura de Goebel es autodual, luego vamos a mostrar un ejemplo relacionados con el diseño de estructuras y finalmente abordamos la extensión de la media proximal.

3.1. Aproximación suave autodual

Una función f es suave en \mathbb{R}^n si f es finita y diferenciable en todo \mathbb{R}^n . En el caso de funciones convexas no diferenciables, se puede hallar una aproximación suave de la función. Vamos a definir una aproximación suave usando la media proximal, lo cual deja en evidencia que se pueden definir otras aproximaciones suaves en términos de la media proximal. El operador $G_\beta f$, es una expresión en términos de la envoltura de Moreau, publicada por Goebel en el año 2009. Esta aproximación es auto dual en el sentido que la conjugada de la aproximación de f es igual a la aproximación de la conjugada de f , lo cual es más sencillo de probar usando la media proximal, para lo cual recordemos que cuando $\mu = 1$ se tiene que

$$\mathcal{P}(f_1, \lambda_1, \dots, f_n, \lambda_n) = (\lambda_1(f_1 + \mathcal{Q})^* + \dots + \lambda_n(f_n + \mathcal{Q})^*)^* - \mathcal{Q}$$

Para $0 \leq \beta \leq 1$ y una función propia semicontinua inferiormente f , se tiene $G_\beta f : \mathbb{R}^n \rightarrow]-\infty, +\infty]$ definida como

$$(3.1) \quad G_\beta f(x) = (1 + \beta)^2 \mathcal{P}(f, \frac{1 - \beta}{1 + \beta}, \mathcal{Q}, \frac{2\beta}{1 + \beta})(\frac{x}{1 + \beta})$$

para todo $x \in \mathbb{R}^n$.

Teorema 3.1. *La aproximación $G_\beta f$ es auto dual, es decir*

$$(G_\beta f)^* = G_\beta f^*$$

Demostración. Vamos a aplicar la conjugada convexa a la aproximación $G_\beta f(x)$, se tiene

$$(G_\beta f(x))^* = \left((1+\beta)^2 \mathcal{P}\left(f, \frac{1-\beta}{1+\beta}, \mathcal{Q}, \frac{2\beta}{1+\beta}\right) \left(\frac{x}{1+\beta}\right) \right)^*$$

Ahora aplicando el Teorema 1.19 y Teorema 1.22 i) tenemos

$$\begin{aligned} (G_\beta f)^* &= \left((1+\beta)^2 \mathcal{P}\left(f, \frac{1-\beta}{1+\beta}, \mathcal{Q}, \frac{2\beta}{1+\beta}\right) \right)^* ((1+\beta)\cdot) \\ &= (1+\beta)^2 \left(\mathcal{P}\left(f, \frac{1-\beta}{1+\beta}, \mathcal{Q}, \frac{2\beta}{1+\beta}\right) \right)^* \left(\frac{(1+\beta)\cdot}{(1+\beta)^2} \right) \\ &= (1+\beta)^2 \mathcal{P}\left(f^*, \frac{1-\beta}{1+\beta}, \mathcal{Q}, \frac{2\beta}{1+\beta}\right) \left(\frac{\cdot}{(1+\beta)} \right) \\ (G_\beta f)^* &= G_\beta f^* \end{aligned}$$

□

Observación 3.1. La envoltura de Moreau no es auto dual, es decir

$$(e_\beta f)^* \neq e_\beta f^*$$

En efecto, pues

$$\begin{aligned} (e_\beta f)^* &= (f \# \beta \star \mathcal{Q})^* \\ &= f^* + (\beta \star \mathcal{Q})^* \\ &= f^* + \beta \mathcal{Q} \\ &= f^* + \beta^{-1} \star \mathcal{Q} \neq e_\beta f^* \\ (e_\beta f)^* &\neq e_\beta f^* \end{aligned}$$

Aproximación de funciones silla

Definición 3.1. Sea K una función silla propia y cerrada, la envoltura de Moreau para $\lambda > 0$ es

$$(3.2) \quad E_\lambda K(x, y) = \inf_{q \in \mathbb{R}^n} \sup_{p \in \mathbb{R}^m} \left\{ K(p, q) - \frac{1}{2\lambda} \|x - p\|^2 + \frac{1}{2\lambda} \|y - q\|^2 \right\}$$

y definimos la envoltura de Goebel para la función silla K con $0 < \lambda < 1$ como

$$(3.3) \quad \mathcal{G}_\lambda K(x, y) = (1 - \lambda^2) E_\lambda K(x) + \frac{\lambda}{2} (\|y\|^2 - \|x\|^2)$$

Dada una función convexa $\phi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ tenemos la envoltura de Goebel en el sentido de (2.4)

$$(3.4) \quad G_\lambda \phi(p, y) = (1 - \lambda^2) \inf_{u \in \mathbb{R}^n, v \in \mathbb{R}^m} \left\{ \phi(u, v) + \frac{1}{2\lambda} \|p - u\|^2 + \frac{1}{2\lambda} \|y - v\|^2 \right\} + \frac{\lambda}{2} \|p\|^2 + \frac{\lambda}{2} \|y\|^2$$

Teorema 3.2. Sea $K : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ una función silla propia y cerrada, $\phi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ es una función propia semicontinua inferiormente y convexa dada por

$$\phi(p, y) = \sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - K(x, y) \}$$

entonces

$$G_\lambda \phi(p, y) = (\mathcal{G}_\lambda K(\cdot, y))^*(p) = \sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - \mathcal{G}_\lambda K(x, y) \}$$

equivalentemente

$$\mathcal{G}_\lambda K(x, y) = (G_\lambda \phi(\cdot, y))^*(x) = \sup_{p \in \mathbb{R}^m} \{ \langle x, p \rangle - G_\lambda \phi(p, y) \}$$

Demostración.

Vea [5] Teorema 3.2

□

3.1.1. Aplicación a la optimización

Consideremos el siguiente problema de optimización

$$(P) \quad \inf \{ f(x) / g_i(x) \leq 0, \ i = 1, 2, \dots, p \}$$

donde $f, g_i : \mathbb{R} \rightarrow \overline{\mathbb{R}}$, sea

$$C = \{ x \in \mathbb{R}^n / g_i \leq 0, \ i = 1, 2, \dots, p \}$$

de modo que el problema (P) equivale a

$$(P) \quad \inf [f + i_C]$$

Lema 3.3. Para cada $x \in \mathbb{R}^n$ se tiene que

$$i_C(x) = \sup_{\lambda \in \mathbb{R}_+^p} \{ \langle \lambda, g(x) \rangle \}$$

Demostración. Se tiene que

$$\sup_{\lambda_i \geq 0} \lambda_i g_i(x) = \begin{cases} 0 & \text{si } g_i(x) \leq 0 \\ +\infty & \text{si } g_i(x) > 0 \end{cases}$$

de modo que

$$i_C(x) = \sup_{\lambda \in \mathbb{R}_+^p} \left\{ \sum_{i=1}^p \lambda_i g_i \right\} = \sup_{\lambda \in \mathbb{R}_+^p} \{ \langle \lambda, g(x) \rangle \}$$

□

La función Lagrangeana de (P) es la función $L : \mathbb{R}^p \times \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ definida

$$L(\lambda, x) = \begin{cases} f(x) + \langle \lambda, g(x) \rangle & \text{si } \lambda \in \mathbb{R}_+^p \\ -\infty & \text{si no} \end{cases}$$

Tenemos que el problema primal es equivalente a

$$(P) \quad \inf_{x \in \mathbb{R}^n} \sup_{\lambda \in \mathbb{R}^p} L(\lambda, x)$$

Definimos el problema dual

$$(D) \quad \sup_{\lambda \in \mathbb{R}^p} \inf_{x \in \mathbb{R}^n} L(\lambda, x)$$

Definimos $\varphi : \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$

$$\varphi(y, x) = \sup_{y^*} \{ \langle y, y^* \rangle + L(y^*, x) \} = (-L(\cdot, x))^*(y)$$

entonces, el problema primal es equivalente a

$$(P) \quad \inf_{x \in \mathbb{R}^n} \varphi(0, x)$$

por otro lado, si $y^* \rightarrow -L(y^*, x)$ es convexa, semicontinua inferiormente y propia, entonces

$$-L(y^*, x) = \sup_{y \in \mathbb{R}^p} \{ \langle y, y^* \rangle - \varphi(y, x) \}$$

luego

$$\inf_{x \in \mathbb{R}^n} L(y^*, x) = - \sup_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^p} \{ \langle y^*, y \rangle - \varphi(y, x) \} = -\varphi^*(y^*, 0)$$

de donde se tiene que

$$(D) \quad \sup_{y^* \in \mathbb{R}^p} -\varphi^*(y^*, 0)$$

Ejemplo 3.1. Consideremos el problema primal (P)

$$\inf_x \phi(x), \text{ donde } \phi(x) = \varphi(0, x)$$

y el respectivo problema dual (D)

$$\sup_y \varphi(y), \text{ donde } \varphi(y) = -\varphi^*(y, 0)$$

aplicando el teorema 3.2, tenemos el problema primal regularizado (P_λ) dado como

$$\inf_x \phi_\lambda(x) = \inf_x G_\lambda \varphi(0, x) = \inf_x G_\lambda \phi(x)$$

y su respectivo problema dual (D_λ) viene dado

$$\sup_y \varphi_\lambda(y) = \sup_y -(s_\lambda \varphi)^*(y, 0) = \sup_y -G_\lambda \varphi^*(y, 0) = \sup_y -G_\lambda(\gamma(y))$$

3.2. Optimización multiobjetivo

En el proceso de fabricar un producto o en el hecho de mejorar la calidad de vida de un individuo se deben tomar decisiones basadas en una gran cantidad de variables con características y comportamiento diferentes y que al evaluarlas adecuadamente elevan notablemente el beneficio final. Ingresamos al campo de la optimización con metodologías tan complejas como el mismo problema.

Las técnicas de optimización adecúan una serie de situaciones de tal manera que se mejoren los resultados por una función objetivo, los problemas reales involucran una serie de objetivos que pueden ser tan relevantes como el que se optimizó, pero también conflictivos que harían inviable la solución encontrada. Prácticamente en cualquier área de las ciencias e ingeniería y en una variedad de contextos se presentan problemas con múltiples objetivos que se contraponen entre sí.

Definición del problema

Estamos interesados en calcular un x tal que

$$\min_x \{f(x)\}$$

Sujeto a

$$\begin{aligned} g_i(x) &\geq 0 && \text{para } i = 1, \dots, n \\ h_j(x) &= 0 && \text{para } j = 1, \dots, p \end{aligned}$$

donde f es un vector de k funciones objetivos $f = (f_1, \dots, f_n)$, x es el vector solución $x = (x_1, \dots, x_r)^t$ de r variables, n es el número de restricciones de desigualdad; y p es el número de restricciones de igualdad.

Pero el problema está en que normalmente no existe un elemento que nos proporcione el óptimo para cada una de las funciones en forma simultánea, esto se debe a la existencia de conflictos entre los objetivos que hacen que la mejora de uno de ellos produzca problemas en alguna de las otras funciones a optimizar, por ejemplo un algoritmo debe ser rápido y dar soluciones de calidad, deseamos un auto que sea rápido, pero consume más combustible, un edificio debe ser económico, seguro y fácil de mantener, deseamos un contenedor de calidad de menor costo y mínimo peso, una estructura debe soportar una carga pero debe tener un diseño de mínimo peso y costo, obviamente tenemos problemas pues para mejorar un objetivo de un problema multiobjetivo, empeoramos otro objetivo, el problema será entonces que los objetivos sean satisfechos en una medida aceptable. Vamos a plantear un método basado en la combinación de objetivos, usaremos una suma ponderada donde se optimizará la función obtenida correspondiente a las distintas funciones objetivo multiplicadas cada una de ellas por un coeficiente de peso, los cuales establecen la importancia relativa de cada objetivo, entonces el problema de optimización multiobjetivo se transforma en otro de optimización escalar de la forma

$$\min_x \{\mathcal{P}_\mu(f, \lambda)(x)\}$$

donde $f = (f_1, \dots, f_n)$, $\lambda = (\lambda_1, \dots, \lambda_n)$ y $\mathcal{P}_\mu(f, \lambda)$ es la media proximal de peso λ con parámetro μ de n funciones f_i con $1 \leq i \leq n$ se define como

$$\mathcal{P}_\mu(f, \lambda) = \lambda_1 \star (f_1 + \mu \star \mathcal{Q}) \# \dots \# \lambda_n \star (f_n + \mu \star \mathcal{Q}) - \mu \star \mathcal{Q}$$

donde $x \in \mathbb{X}$, tenemos

$$\mathcal{P}_\mu(f, \lambda)(x) = \lambda_1(f_1 + \mu \star \mathcal{Q})\left(\frac{x_1}{\lambda_1}\right) \# \dots \# \lambda_n(f_n + \mu \star \mathcal{Q})\left(\frac{x_n}{\lambda_n}\right) - \mu \mathcal{Q}\left(\frac{x}{\mu}\right)$$

este planteamiento tiene la ventaja que los dominios de las distintas funciones f_i pueden ser disjuntos y además la función $\mathcal{P}_\mu(f, \lambda)$ tiene la propiedad autodual.

Ejemplo 3.2. El costo y el peso son problemas para encontrar el diseño óptimo de un contenedor, formulándose matemáticamente de la siguiente manera

$$\begin{aligned} \text{Costo } f_1 &= 1000x_1^2 + 2000x_1x_2 \\ \text{Peso } f_2 &= 4x_1^2 + 32x_1x_2 \end{aligned}$$

donde: x_1 = dimensión de los lados de la base del contenedor y x_2 = dimensión de la altura del contenedor.

El problema consiste en fabricar contenedores de base cuadrada, abiertos por la parte de arriba con una capacidad o volumen de $10m^3$. El fabricante desea minimizar el coste de los contenedores, así como minimizar su peso. Ahora bien, como los materiales más ligeros son los más caros nos encontramos ante un claro conflicto entre objetivos. En concreto el material con el que se va a construir el fondo del contenedor cuesta $1000s/m^2$ y pesa $4Kg/m^2$, mientras que el material con el que se van a construir las caras del contenedor cuestan $500s/m^2$ y pesa $8Kg/m^2$. Si representamos por x_1 la dimensión de los lados de la base del contenedor y por x_2 la dimensión de su altura, tenemos el siguiente problema multiobjetivo planteado en [11]

$$\begin{aligned} &\min\{f_1, f_2\} \\ &f_1(x) = 1000x_1^2 + 2000x_1x_2 \\ &f_2(x) = 4x_1^2 + 32x_1x_2 \\ &\text{sujeto a} \\ &x_1^2x_2 = 10 \end{aligned}$$

Para resolver este problema usaremos la relación 2.3 con $\mu = 1$, es decir

$$(3.5) \quad \mathcal{P}(f, \lambda)(x) = \inf_{\sum_{i=1}^n x_i = x} \sum_{i=1}^n \left\{ \lambda_i \left[f_i\left(\frac{x_i}{\lambda_i}\right) + \mathcal{Q}\left(\frac{x_i}{\lambda_i}\right) \right] \right\} - \mathcal{Q}(x)$$

En nuestro caso debido a que tenemos dos funciones $f_1(x)$ y $f_2(x)$, planteamos la ecuación 3.5 de la forma siguiente

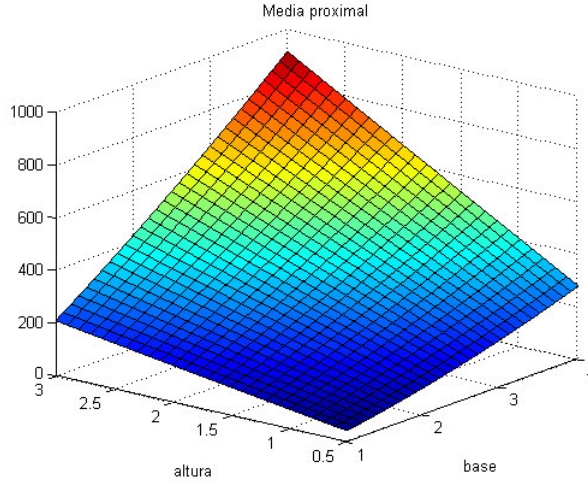


Figura 3.1: Media proximal de f_1 y f_2 tomando $\lambda = 0,5$

$$\mathcal{P}(f, \lambda)(p) = \inf_x \left\{ (1 - \lambda) \left[f_1\left(\frac{x}{1 - \lambda}\right) + \mathcal{Q}\left(\frac{x}{1 - \lambda}\right) \right] + \lambda \left[f_2\left(\frac{p - x}{\lambda}\right) + \mathcal{Q}\left(\frac{p - x}{\lambda}\right) \right] \right\} - \mathcal{Q}(p)$$

se tiene

(3.6)

$$\begin{aligned} \mathcal{P}(f, \lambda)(p) = \inf_x \{ & (1 - \lambda) \left[1000\left(\frac{x_1}{1 - \lambda}\right)^2 + \frac{2000x_1x_2}{(1 - \lambda)^2} + \mathcal{Q}\left(\frac{x_1}{1 - \lambda}, \frac{x_2}{1 - \lambda}\right) \right] + \\ & + \lambda \left[4\left(\frac{p_1 - x_1}{\lambda}\right)^2 + \frac{32(p_1 - x_1)(p_2 - x_2)}{\lambda^2} + \mathcal{Q}\left(\frac{p_1 - x_1}{\lambda}, \frac{p_2 - x_2}{\lambda}\right) \right] \} - \mathcal{Q}(p) \end{aligned}$$

Definimos la función

$$\begin{aligned} h(x) = (1 - \lambda) \left[1000\left(\frac{x_1}{1 - \lambda}\right)^2 + \frac{2000x_1x_2}{(1 - \lambda)^2} + \mathcal{Q}\left(\frac{x_1}{1 - \lambda}, \frac{x_2}{1 - \lambda}\right) \right] + \\ + \lambda \left[4\left(\frac{p_1 - x_1}{\lambda}\right)^2 + \frac{32(p_1 - x_1)(p_2 - x_2)}{\lambda^2} + \mathcal{Q}\left(\frac{p_1 - x_1}{\lambda}, \frac{p_2 - x_2}{\lambda}\right) \right] \end{aligned}$$

y vamos a obtener su mínimo valor, para lo cual calculamos las derivadas parciales de la función h respecto de x_1 y de x_2

$$\begin{aligned} \frac{\partial h(x_1, x_2)}{\partial x_1} &= -\lambda \left(\frac{p_1 - x_1}{\lambda} + \frac{8(p_1 - x_1)}{\lambda^2} + \frac{32(p_2 - x_2)}{\lambda^2} \right) - \\ &\quad - (\lambda - 1) \left(\frac{2000x_1}{(\lambda - 1)^2} + \frac{2000x_2}{(\lambda - 1)^2} - \frac{x_1}{x_1 - 1} \right) \\ \frac{\partial h(x_1, x_2)}{\partial x_2} &= -\lambda \left(\frac{p_2 - x_2}{\lambda} + \frac{32p_1 - 32x_1}{\lambda^2} \right) - \left(\frac{2000x_1}{(\lambda - 1)^2} - \frac{x_2}{\lambda - 1} \right) (\lambda - 1) \end{aligned}$$

ahora vamos a resolver el sistema

$$\begin{cases} \frac{\partial h(x_1, x_2)}{\partial x_1} = 0 \\ \frac{\partial h(x_1, x_2)}{\partial x_2} = 0 \end{cases}$$

y obtenemos los valores

$$x_1 = -\frac{(\lambda - 1)(512p_1 - 31480\lambda p_1 - 16\lambda p_2 + 7\lambda^2 p_1 - 1016\lambda^2 p_2 + \lambda^3 p_1)}{2(-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256)}$$

$$x_2 = -\frac{(\lambda - 1)(512p_2 - 24016\lambda p_1 - 31484\lambda p_2 + 1016\lambda^2 p_1 - 997\lambda^2 p_2 + \lambda^3 p_2)}{2(-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256)}$$

Reemplazando en la relación 3.6 obtenemos la media proximal de las funciones f_1 y f_2 dada como

$$\begin{aligned} \mathcal{P}(f, \lambda)(p) = & \frac{\lambda^2(261816p_1^2 + 19516p_1p_2 + 258064p_2^2) + 512000p_1p_2}{-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256} + \\ & + \frac{\lambda(3731871p_1^2 + 31467992p_1p_2 - 8128p_2^2) + 256064p_1^2 + 64p_2^2}{-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256} + \\ & + \frac{\lambda^3(249p_1^2 + 492p_2p_1)}{-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256} + \frac{3p_1^2}{4} + \frac{3p_2^2}{4} \end{aligned}$$

Vamos a ordenar respecto de p_1 y p_2

$$\begin{aligned} \mathcal{P}(f, \lambda)(p) = & p_1^2 \left(\frac{249\lambda^3 + 261816\lambda^2 + 3731871\lambda + 256064}{-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256} + \frac{3}{4} \right) + \\ & + p_1p_2 \left(\frac{492\lambda^3 + 19516\lambda^2 + 31467992\lambda + 512000}{-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256} \right) + \\ & + p_2^2 \left(\frac{258064\lambda^2 - 8128\lambda + 64}{-\lambda^4 + 998\lambda^3 + 967263\lambda^2 + 31484\lambda + 256} + \frac{3}{4} \right) \end{aligned}$$

Ahora estamos interesados en calcular un x tal que resuelva el siguiente problema de optimización escalar

$$(3.7) \quad \begin{aligned} & \min_x \{ \mathcal{P}(f, \lambda)(x) \} \\ & \text{sujeto a} \\ & x_1^2 x_2 = 10 \end{aligned}$$

Para lograr nuestro objetivo vamos a utilizar el siguiente código de Matlab

```
clear;clc;
landa=linspace(0,1,100);
resultado=[0      0      0      ];
```

```

for i=2:99
lan=landa(i);
fun=@(x)((lan^2*(261816*x(1)^2 + 19516*x(1)*x(2) + 258064*x(2)^2) +...
512000*x(1)*x(2) + lan*(3731871*x(1)^2 + 31467992*x(1)*x(2) -...
8128*x(2)^2) + 256064*x(1)^2 + 64*x(2)^2 + lan^3*(249*x(1)^2 + ...
492*x(2)*x(1)))./((- lan^4 + 998*lan^3 + 967263*lan^2 + ...
31484*lan + 256) + (3*x(1)^2)/4 + (3*x(2)^2)/4 );
p0=[1; 1];
A=[];b=[];Aeq=[];beq=[]; lb=[0.5 0.5];ub=[4 4];
[q,mprox] = fmincon(fun,p0,A,b,Aeq,beq,lb,ub,@restricn1)
resultado=[resultado; lan      q'   ];
end
syms a b;
f1=1000*a^2+2000*a*b;
f2=4*a^2+32*a*b;
re=[0 0 0];
for i=2:99
costo=subs( subs(f1,a,resultado(i,2)),b,resultado(i,3) );
peso=subs( subs(f2,a,resultado(i,2)),b,resultado(i,3) );
re=[re; mprox(i)  costo  peso ];
end
format short e;  clc;
disp('')
disp('-----')
disp(' lambda  Base  Altura  M. Proximal  Costo  Peso  ')
disp('-----')
disp('')
tabla=[resultado  re];
disp(tabla)

```

este código nos permite obtener los siguientes resultados para diferentes pesos o valores de $\lambda \in \langle 0; 1 \rangle$

lambda	Base	Altura	M. Proximal	Costo	Peso
1.01e-02	2.4172e+00	1.7114e+00	7.6618e+03	1.4117e+04	1.5575e+02
2.02e-02	2.5859e+00	1.4955e+00	5.1788e+03	1.4421e+04	1.5050e+02
3.03e-02	2.7053e+00	1.3664e+00	3.8772e+03	1.4711e+04	1.4756e+02
4.04e-02	2.7946e+00	1.2804e+00	3.0847e+03	1.4967e+04	1.4574e+02
5.05e-02	2.8642e+00	1.2190e+00	2.5549e+03	1.5186e+04	1.4454e+02
6.06e-02	2.9198e+00	1.1730e+00	2.1773e+03	1.5375e+04	1.4370e+02
7.07e-02	2.9653e+00	1.1373e+00	1.8953e+03	1.5538e+04	1.4309e+02
8.08e-02	3.0031e+00	1.1089e+00	1.6770e+03	1.5678e+04	1.4263e+02
9.09e-02	3.0349e+00	1.0857e+00	1.5032e+03	1.5801e+04	1.4228e+02
1.01e-01	3.0620e+00	1.0666e+00	1.3618e+03	1.5908e+04	1.4201e+02

1.11e-01	3.0853e+00	1.0505e+00	1.2446e+03	1.6002e+04	1.4179e+02
1.21e-01	3.1055e+00	1.0369e+00	1.1458e+03	1.6084e+04	1.4162e+02
1.31e-01	3.1232e+00	1.0252e+00	1.0615e+03	1.6158e+04	1.4148e+02
1.41e-01	3.1387e+00	1.0151e+00	9.8876e+02	1.6223e+04	1.4136e+02
1.51e-01	3.1524e+00	1.0063e+00	9.2536e+02	1.6282e+04	1.4126e+02
1.61e-01	3.1645e+00	9.9862e-01	8.6961e+02	1.6334e+04	1.4118e+02
1.71e-01	3.1752e+00	9.9186e-01	8.2022e+02	1.6381e+04	1.4111e+02
1.81e-01	3.1848e+00	9.8588e-01	7.7618e+02	1.6423e+04	1.4105e+02
1.91e-01	3.1934e+00	9.8058e-01	7.3665e+02	1.6461e+04	1.4100e+02
2.02e-01	3.2011e+00	9.7586e-01	7.0098e+02	1.6495e+04	1.4095e+02
2.12e-01	3.2081e+00	9.7166e-01	6.6864e+02	1.6526e+04	1.4092e+02
2.22e-01	3.2143e+00	9.6790e-01	6.3919e+02	1.6554e+04	1.4088e+02
2.32e-01	3.2199e+00	9.6453e-01	6.1225e+02	1.6579e+04	1.4085e+02
2.42e-01	3.2250e+00	9.6151e-01	5.8752e+02	1.6602e+04	1.4083e+02
2.52e-01	3.2295e+00	9.5880e-01	5.6473e+02	1.6623e+04	1.4081e+02
2.62e-01	3.2336e+00	9.5636e-01	5.4368e+02	1.6641e+04	1.4079e+02
2.72e-01	3.2373e+00	9.5418e-01	5.2416e+02	1.6658e+04	1.4077e+02
2.82e-01	3.2407e+00	9.5221e-01	5.0602e+02	1.6673e+04	1.4075e+02
2.92e-01	3.2437e+00	9.5045e-01	4.8912e+02	1.6687e+04	1.4074e+02
3.03e-01	3.2464e+00	9.4887e-01	4.7333e+02	1.6700e+04	1.4073e+02
3.13e-01	3.2488e+00	9.4746e-01	4.5855e+02	1.6711e+04	1.4072e+02
3.23e-01	3.2509e+00	9.4620e-01	4.4469e+02	1.6721e+04	1.4071e+02
3.33e-01	3.2529e+00	9.4507e-01	4.3165e+02	1.6730e+04	1.4070e+02
3.43e-01	3.2546e+00	9.4407e-01	4.1938e+02	1.6738e+04	1.4069e+02
3.53e-01	3.2561e+00	9.4319e-01	4.0781e+02	1.6745e+04	1.4069e+02
3.63e-01	3.2575e+00	9.4242e-01	3.9687e+02	1.6751e+04	1.4068e+02
3.73e-01	3.2586e+00	9.4174e-01	3.8652e+02	1.6756e+04	1.4068e+02
3.83e-01	3.2596e+00	9.4116e-01	3.7670e+02	1.6761e+04	1.4067e+02
3.93e-01	3.2605e+00	9.4065e-01	3.6739e+02	1.6765e+04	1.4067e+02
4.04e-01	3.2612e+00	9.4023e-01	3.5855e+02	1.6768e+04	1.4066e+02
4.14e-01	3.2618e+00	9.3988e-01	3.5013e+02	1.6771e+04	1.4066e+02
4.24e-01	3.2623e+00	9.3960e-01	3.4211e+02	1.6773e+04	1.4066e+02
4.34e-01	3.2627e+00	9.3938e-01	3.3446e+02	1.6775e+04	1.4066e+02
4.44e-01	3.2630e+00	9.3922e-01	3.2715e+02	1.6776e+04	1.4066e+02
4.54e-01	3.2632e+00	9.3911e-01	3.2017e+02	1.6777e+04	1.4066e+02
4.64e-01	3.2633e+00	9.3906e-01	3.1349e+02	1.6778e+04	1.4066e+02
4.74e-01	3.2633e+00	9.3905e-01	3.0710e+02	1.6778e+04	1.4066e+02
4.84e-01	3.2632e+00	9.3909e-01	3.0097e+02	1.6778e+04	1.4066e+02
4.94e-01	3.2631e+00	9.3917e-01	2.9508e+02	1.6777e+04	1.4066e+02
5.05e-01	3.2629e+00	9.3929e-01	2.8944e+02	1.6776e+04	1.4066e+02
5.15e-01	3.2626e+00	9.3945e-01	2.8401e+02	1.6775e+04	1.4066e+02
5.25e-01	3.2623e+00	9.3964e-01	2.7879e+02	1.6773e+04	1.4066e+02
5.35e-01	3.2619e+00	9.3987e-01	2.7377e+02	1.6771e+04	1.4066e+02
5.45e-01	3.2614e+00	9.4012e-01	2.6893e+02	1.6769e+04	1.4066e+02
5.55e-01	3.2609e+00	9.4041e-01	2.6427e+02	1.6767e+04	1.4067e+02
5.65e-01	3.2604e+00	9.4072e-01	2.5977e+02	1.6764e+04	1.4067e+02
5.75e-01	3.2598e+00	9.4106e-01	2.5543e+02	1.6762e+04	1.4067e+02

5.85e-01	3.2592e+00	9.4143e-01	2.5124e+02	1.6759e+04	1.4067e+02
5.95e-01	3.2585e+00	9.4182e-01	2.4719e+02	1.6756e+04	1.4068e+02
6.06e-01	3.2578e+00	9.4223e-01	2.4328e+02	1.6752e+04	1.4068e+02
6.16e-01	3.2570e+00	9.4266e-01	2.3949e+02	1.6749e+04	1.4068e+02
6.26e-01	3.2562e+00	9.4312e-01	2.3583e+02	1.6745e+04	1.4069e+02
6.36e-01	3.2554e+00	9.4359e-01	2.3228e+02	1.6741e+04	1.4069e+02
6.46e-01	3.2546e+00	9.4408e-01	2.2884e+02	1.6738e+04	1.4069e+02
6.56e-01	3.2537e+00	9.4459e-01	2.2551e+02	1.6733e+04	1.4070e+02
6.66e-01	3.2528e+00	9.4511e-01	2.2228e+02	1.6729e+04	1.4070e+02
6.76e-01	3.2519e+00	9.4565e-01	2.1914e+02	1.6725e+04	1.4070e+02
6.86e-01	3.2509e+00	9.4620e-01	2.1610e+02	1.6721e+04	1.4071e+02
6.96e-01	3.2500e+00	9.4677e-01	2.1314e+02	1.6716e+04	1.4071e+02
7.07e-01	3.2490e+00	9.4735e-01	2.1027e+02	1.6712e+04	1.4072e+02
7.17e-01	3.2479e+00	9.4795e-01	2.0748e+02	1.6707e+04	1.4072e+02
7.27e-01	3.2469e+00	9.4855e-01	2.0476e+02	1.6702e+04	1.4072e+02
7.37e-01	3.2458e+00	9.4917e-01	2.0212e+02	1.6697e+04	1.4073e+02
7.47e-01	3.2448e+00	9.4980e-01	1.9955e+02	1.6692e+04	1.4073e+02
7.57e-01	3.2437e+00	9.5044e-01	1.9705e+02	1.6687e+04	1.4074e+02
7.67e-01	3.2426e+00	9.5109e-01	1.9462e+02	1.6682e+04	1.4074e+02
7.77e-01	3.2414e+00	9.5175e-01	1.9225e+02	1.6677e+04	1.4075e+02
7.87e-01	3.2403e+00	9.5242e-01	1.8993e+02	1.6672e+04	1.4075e+02
7.97e-01	3.2392e+00	9.5310e-01	1.8768e+02	1.6667e+04	1.4076e+02
8.08e-01	3.2380e+00	9.5378e-01	1.8548e+02	1.6661e+04	1.4077e+02
8.18e-01	3.2368e+00	9.5447e-01	1.8334e+02	1.6656e+04	1.4077e+02
8.28e-01	3.2356e+00	9.5518e-01	1.8125e+02	1.6650e+04	1.4078e+02
8.38e-01	3.2344e+00	9.5588e-01	1.7921e+02	1.6645e+04	1.4078e+02
8.48e-01	3.2332e+00	9.5660e-01	1.7722e+02	1.6639e+04	1.4079e+02
8.58e-01	3.2320e+00	9.5732e-01	1.7527e+02	1.6634e+04	1.4079e+02
8.68e-01	3.2308e+00	9.5805e-01	1.7337e+02	1.6628e+04	1.4080e+02
8.78e-01	3.2295e+00	9.5878e-01	1.7152e+02	1.6623e+04	1.4081e+02
8.88e-01	3.2283e+00	9.5952e-01	1.6970e+02	1.6617e+04	1.4081e+02
8.98e-01	3.2270e+00	9.6027e-01	1.6793e+02	1.6611e+04	1.4082e+02
9.09e-01	3.2258e+00	9.6102e-01	1.6619e+02	1.6606e+04	1.4082e+02
9.19e-01	3.2245e+00	9.6178e-01	1.6450e+02	1.6600e+04	1.4083e+02
9.29e-01	3.2232e+00	9.6254e-01	1.6284e+02	1.6594e+04	1.4084e+02
9.39e-01	3.2219e+00	9.6330e-01	1.6121e+02	1.6588e+04	1.4084e+02
9.49e-01	3.2207e+00	9.6407e-01	1.5962e+02	1.6583e+04	1.4085e+02
9.59e-01	3.2194e+00	9.6485e-01	1.5806e+02	1.6577e+04	1.4086e+02
9.69e-01	3.2181e+00	9.6563e-01	1.5654e+02	1.6571e+04	1.4086e+02
9.79e-01	3.2168e+00	9.6641e-01	1.5505e+02	1.6565e+04	1.4087e+02
9.89e-01	3.2155e+00	9.6719e-01	1.5358e+02	1.6559e+04	1.4088e+02

Observamos los resultados obtenidos en [11] mostrados en la siguiente tabla, notamos que nuestros resultados fueron satisfactorios. Este ejemplo sugiere que en situaciones reales, podremos probar con diferentes valores para λ y estos resultados nos permitirán tener un criterio aceptable para una la solución, es decir nos permitirán tomar una decisión adecuada a nuestros requerimientos.

Base del contador - m - x_1	Altura del contenedor - m - x_2	Coste - ptas - $f_1(x)$	Peso - Kg - $f_2(x)$
3,42	0,85	17752	140
3,41	0,86	17500	140,35
3,31	0,91	17000	140,50
3,20	0,97	16500	140,94
3,08	1,05	16000	141,79
2,95	1,15	15500	143,21
2,81	1,27	15000	145,54
2,62	1,46	14500	149,57
2,31	1,86	14000	159,60
2,15	2,15	13925	167

Figura 3.2: Resultados obtenidos de [11]

3.3. Extensión de la media proximal

En esta sección vamos a considerar a $\lambda_1, \dots, \lambda_n$ como números reales no negativos tales que $\sum_{i=1}^n \lambda_i = 1$. Se plantea la pregunta. ¿Puede trabajar la media proximal si las funciones f_i no son convexas ?.

En los textos [9] y [10], se dan muchas y muy interesantes proposiciones de las cuales hemos extraído algunas de ellas y las presentamos a continuación, las cuales motivan una extensión de la idea de la media proximal paramétrica a otra media proximal que si puede trabajar con funciones no necesariamente convexas.

Teorema 3.4. Relación de la conjugada y la envoltura proximal

Sean $f_i : \mathbb{R} \rightarrow \langle -\infty, +\infty \rangle$, $i = 1, \dots, n$, funciones propias semicontinuas inferiormente (sci), no necesariamente convexas, $f = (f_1, \dots, f_n)$, $\lambda = (\lambda_1, \dots, \lambda_n)$, se cumple

- i) La envoltura de Moreau se puede relacionar con la envoltura proximal del modo siguiente :

$$(f + r\mathcal{Q})^*(w) = (-e_{r^{-1}}f + r\mathcal{Q})(r^{-1}w)$$

ii) $\mathcal{P}_r(f, \lambda) = \mathcal{P}_r(f_1, \lambda_1, \dots, f_n, \lambda_n) = -e_r(-\lambda_1 e_r f_1 - \lambda_2 e_r f_2 - \dots - \lambda_n e_r f_n)$

iii) Si $\lambda > 0$ se cumple $h_\lambda f(x) = -e_r(-e_r f(x)) = (f + \lambda^{-1}\mathcal{Q})^{**}(x) - \lambda^{-1}\mathcal{Q}(x)$

iv) $-e_r(-e_r f) = f$ si y solo si f es sci y $f + r\mathcal{Q}$ es convexo .

Demostración.

i) Evaluando tenemos que

$$\begin{aligned}
(f + r\mathcal{Q})^*(w) &= \sup_x \{ \langle x, w \rangle - f(x) - r\mathcal{Q}(x) \} \\
&\quad -|x|^2 \leq -|x - r^{-1}w|^2 + |r^{-1}w|^2 - \langle x, r^{-1}w \rangle \\
&\quad -r\mathcal{Q}(x) \leq -r\mathcal{Q}(x - r^{-1}w) + r^{-1}\mathcal{Q}(w) - \langle x, w \rangle \\
&= \sup_x \{ \langle x, w \rangle - f(x) - r\mathcal{Q}(x - r^{-1}w) + r^{-1}\mathcal{Q}(w) - \langle x, w \rangle \} \\
&= \sup_x \{ r^{-1}\mathcal{Q}(w) - f(x) - r\mathcal{Q}(x - r^{-1}w) \} \\
&= r^{-1}\mathcal{Q}(w) + \sup_x \{ -f(x) - r\mathcal{Q}(x - r^{-1}w) \} \\
&= r^{-1}\mathcal{Q}(w) - \inf_x \{ f(x) + r\mathcal{Q}(x - r^{-1}w) \} \\
&= r^{-1}\mathcal{Q}(w) - (f \sharp r\mathcal{Q})(r^{-1}w) \\
&= r\mathcal{Q}(r^{-1}w) - (f \sharp r^{-1} \star \mathcal{Q})(r^{-1}w) \\
&= r\mathcal{Q}(r^{-1}w) - e_{r^{-1}}f(r^{-1}w) \\
(f + r\mathcal{Q})^*(w) &= (r\mathcal{Q} - e_{r^{-1}}f)(r^{-1}w)
\end{aligned}$$

ii) De la definición de media proximal con parámetro r tenemos

$$\begin{aligned}
\mathcal{P}_r(f, \lambda) &= (\lambda_1(f_1 + r \star \mathcal{Q})^* + \dots + \lambda_n(f_n + r \star \mathcal{Q})^*)^* - r \star \mathcal{Q} \\
&= (\lambda_1(f_1 + r^{-1}\mathcal{Q})^* + \dots + \lambda_n(f_n + r^{-1}\mathcal{Q})^*)^* - r \star \mathcal{Q} \\
&= (\lambda_1(r^{-1}\mathcal{Q} - e_r f_1)^*(rx) + \dots + \lambda_n(r^{-1}\mathcal{Q} - e_r f_n)^*(rx))^* - r \star \mathcal{Q} \\
&= (r\mathcal{Q}(x) - e_r(-\lambda_1 e_r f_1 - \dots - \lambda_n e_r f_n)(rx))^* - r \star \mathcal{Q} \\
&= (r\mathcal{Q} - e_{r^{-1}}(-\lambda_1 e_r f_1 - \dots - \lambda_n e_r f_n)(rx))(r^{-1}x) - r \star \mathcal{Q} \\
&= \not{r}^{-1}\mathcal{Q}(x) - (e_{r^{-1}}(-\lambda_1 e_r f_1 - \dots - \lambda_n e_r f_n)(rx)) - \not{r} \star \mathcal{Q} \\
&= -([e_r(-\lambda_1 e_r f_1 - \dots - \lambda_n e_r f_n)(rx)])(r^{-1}x) \\
&= -e_r(-\lambda_1 e_r f_1 - \dots - \lambda_n e_r f_n)
\end{aligned}$$

iii) Para esta parte de la demostración usamos el Teorema 2.8, tenemos que

$$\begin{aligned}
h_\lambda f(x) &= \sup_w \{ e_\lambda f(w) - \frac{1}{2\lambda} |x - w|^2 \} \\
&= \sup_w \{ \langle x, w \rangle + e_\lambda f(w) - \lambda^{-1}\mathcal{Q}(w) - \lambda^{-1}\mathcal{Q}(x) \} \\
&= \sup_w \{ \langle x, w \rangle + e_\lambda f(w) - \lambda^{-1}\mathcal{Q}(w) \} - \lambda^{-1}\mathcal{Q}(x) \\
&= \sup_w \{ \langle x, w \rangle + (f + \lambda^{-1}\mathcal{Q})^*(w) \} - \lambda^{-1}\mathcal{Q}(x) \\
&= (f + \lambda^{-1}\mathcal{Q})^{**}(x) - \lambda^{-1}\mathcal{Q}(x)
\end{aligned}$$

iv) Para culminar con la demostración de este teorema, planteamos a partir de las hipótesis que $(f + r^{-1}\mathcal{Q})^{**} = (f + r^{-1}\mathcal{Q})$ y así se tiene que usando la proposición anterior $-e_r(-e_r f)(x) = (f + \lambda^{-1}\mathcal{Q})^{**}(x) - \lambda^{-1}\mathcal{Q}(x) = (f + \lambda^{-1}\mathcal{Q})(x) - \lambda^{-1}\mathcal{Q}(x) = f(x)$ si y solo si f es s.c.i. y $f + r\mathcal{Q}$ es convexo.

□

Definición 3.2. Sean $f_i : \mathbb{R} \rightarrow \langle -\infty, +\infty \rangle, i = 1, \dots, n$ funciones propias semicontinua inferiormente (sci), no necesariamente convexas, $f = (f_1, \dots, f_n)$, $\lambda = (\lambda_1, \dots, \lambda_n)$ definimos la *media proximal para funciones no convexas* como

$$\mathcal{P}_r(f, \lambda) = \mathcal{P}_r(f_1, \lambda_1, f_2, \lambda_2, \dots, f_n, \lambda_n) = -e_r(-\lambda_1 e_r f_1 - \lambda_2 e_r f_2 - \dots - \lambda_n e_r f_n)$$

donde si $\lambda_i = 1$ se tiene:

$$\mathcal{P}_r(f, \lambda, \lambda_i) = f_i \text{ si } f_i + r\mathcal{Q} \text{ es convexo}$$

Ejemplo 3.3. Sean las funciones

$$f_1(x) = \begin{cases} -|x| & \text{para } x \leq 0, \\ 0 & \text{para } 0 < x \end{cases} \quad \text{y } f_2(x) = -|x| + 20$$

no son convexas, pero las funciones $f_1 + 40\mathcal{Q}$ y $f_2 + 40\mathcal{Q}$ resultan ser convexas lo cual se puede apreciar en los gráficos siguientes, por lo tanto tenemos las funciones f_1 y f_2 que cumplen el Teorema 3.4 y aquí aplicamos la Definición 3.2, así tenemos las funciones $f_i : \mathbb{R} \rightarrow \langle -\infty, +\infty \rangle, i = 1, 2$ propias semicontinuas inferiormente (s.c.i.), no necesariamente convexas, $f = (f_1, f_2)$, $\lambda = (\lambda_1, \lambda_2)$, tenemos la media proximal para funciones no convexas, dada como

$$\mathcal{P}_r(f, \lambda) = \mathcal{P}_r(f_1, \lambda_1, f_2, \lambda_2) = -e_r(-\lambda_1 e_r f_1 - \lambda_2 e_r f_2)$$

donde si $\lambda_i = 1$ se tiene:

$$\mathcal{P}_r(f, \lambda, \lambda_i) = f_i \text{ pues } f_i + r\mathcal{Q} \text{ es convexo}$$

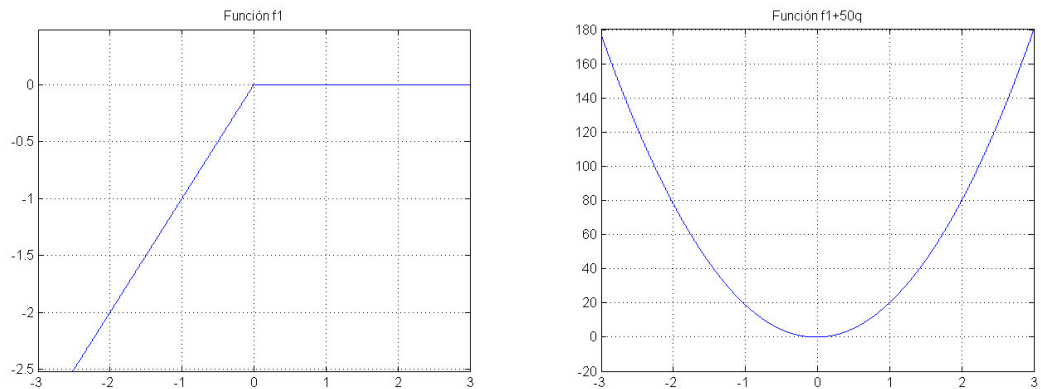


Figura 3.3: Funciones f_1 y $f_1 + 40\mathcal{Q}$ la cual es convexa

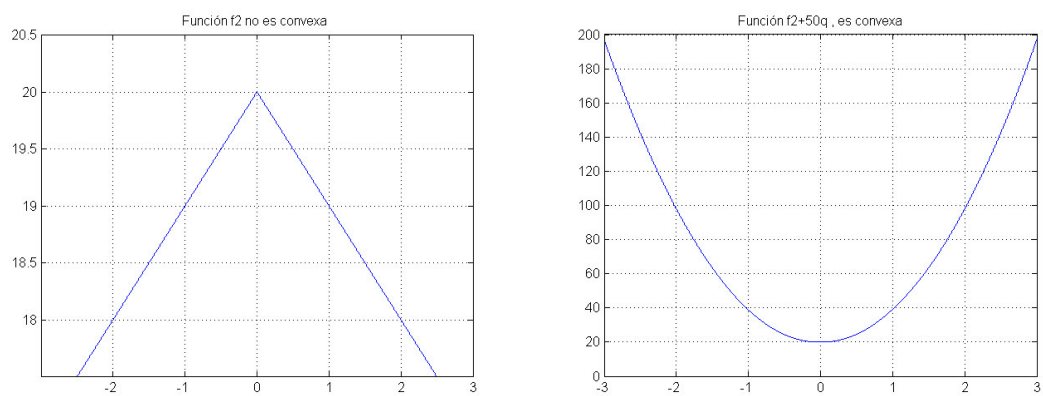


Figura 3.4: Funciones f_2 y $f_2 + 40Q$ la cual es convexa

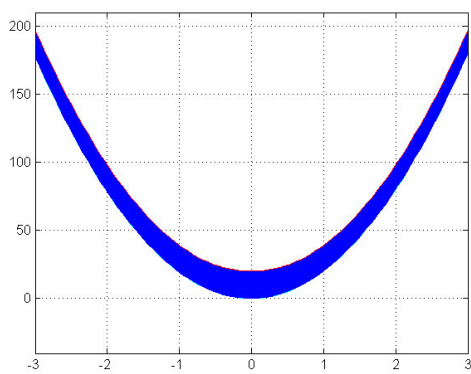


Figura 3.5: Media proximal de f_1 y f_2

Bibliografía

- [1] Hedy Attouch, Giuseppe Buttazzo, Gerard Michaille. *Variational Analysis in Sobolev and B Spaces Applications to PDEs and Optimization*. SIAM 2006.
- [2] Leonard D. Berkovitz. *Convexity and Optimization in \mathbb{R}^n* John Wiley. 2002.
- [3] Jonathan M. Borwein and Jon Vanderwerff, *Convex Functions: Constructions, Characterizations and Counterexamples*, CMS Books in Mathematics, 2nd edition, Springer-Verlag, New York, 2008.
- [4] Boyd, S. and Vandenberghe, L., *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.
- [5] Rafal Goebel. *Self-dual smoothing of convex and saddle functions*, J. Convex Analysis, 2008, 179-190.
- [6] R. Goebel. *Convex Optimal Control Problems with Smooth Hamiltonians* SIAM J. control optim. 43(5):1781-1811, 2005.
- [7] H. Bauschke, R. Goebel, Y. Lucet and X. Wang, *The proximal average: basic theory*, "SIAM Journal on Optimization, vol 19(2), pp. 766-785, 2009.
- [8] R. Ellaia J.B Hiriart Urruty, *The conjugate of the difference of functions*, Journal of optimization theory and applications, vol 49(3), pp.493-498, 1986.
- [9] Rockafellar, R.T., *Convex Analysis*, Princeton University Press, 1970.
- [10] Rockafellar, R.T. Roger J.B Wets, *Variational Analysis*, Springer-Verlag, 1997.
- [11] Romero Carlos, *Análisis de las decisiones multicriterio*, Idefe, publicaciones de ingeniería de sistemas, Madrid 1996.
- [12] C. Zalinescu, *Convex analysis in general vector spaces*, World Scientific Publishing Co. Pte. Ltd. 2002.